

UNIVERSITÉ DE SHERBROOKE
Faculté de génie
Département de génie électrique et de génie informatique

INTÉGRATION ET ÉVALUATION DE CAPACITÉS INTERACTIVES D'UN ROBOT HUMANOÏDE

Mémoire de maîtrise
Spécialité : génie électrique

Vincent ROUSSEAU

Jury : François MICHAUD (directeur)
Charles-Antoine BRUNET
Ruben GONZALEZ-RUBIO

Sherbrooke (Québec) Canada

Septembre 2011

RÉSUMÉ

Le domaine de l'Interaction Humain-Robot (HRI) est en pleine expansion. En effet, de plus en plus de plateformes robotiques sont mises en œuvre pour faire évoluer ce domaine. Sur ces plateformes, toujours plus de modalités d'interaction sont mises en place telles que les mouvements corporels, la reconnaissance de gestes ou d'objets, la reconnaissance et la synthèse vocale ou encore la mobilité, pour pouvoir effectuer l'interaction la plus complète et la plus naturelle pour l'humain. Mais ceci amène aussi une complexité croissante de l'intégration de ces modalités sur une seule et même plateforme.

Aussi, le domaine HRI étant à ses débuts, la méthodologie expérimentale des travaux se limite le plus souvent à des preuves de concept éprouvées en laboratoire ou en milieux ouverts non contrôlés. Il se trouve que peu de chercheurs présentent une démarche structurée et rigoureuse pour l'évaluation expérimentale d'interaction humain-robot en milieux ouverts, et il en résulte des recherches de types exploratoires qui examinent principalement la complexité technologique des modalités interactives à mettre en œuvre, et non l'impact de ces modalités sur la qualité des interactions.

Le but de l'étude présentée dans ce document est d'étudier l'intégration de plusieurs modalités interactives sur un robot mobile humanoïde telles que la parole, les gestes et la mobilité sur la qualité des interactions humain-robot. Plus spécifiquement, le contexte de l'étude consiste à examiner l'impact de modalités interactives sur la capacité du robot à attirer l'attention d'une personne et à engager une interaction avec elle. Le scénario expérimental consiste à permettre au robot, à partir de la parole, d'expressions faciales, de mouvement de la tête, de gestes avec son bras et de sa mobilité, de demander de l'assistance à une personne à proximité de lui remettre un objet se trouvant au sol. L'hypothèse sous-jacente est que l'intégration de l'ensemble de ces modalités devrait améliorer la capacité du robot à engager des personnes à interagir avec lui.

Des expérimentations ont été faites en milieu contrôlé et non-contrôlé selon deux protocoles expérimentaux : une étude des modalités à l'intérieur d'une population, et une étude de variation entre individus. D'une manière générale, il en ressort que l'ajout de modalités améliore la qualité de l'engagement de l'interaction par le robot, mais qu'il faut porter une attention particulière à l'approche de la personne par le robot, principalement pour les personnes non familières avec ce dernier. De plus, les observations indiquent qu'il est plus facile d'obtenir des résultats significatifs en environnement contrôlé, elles permettent d'identifier des pistes d'amélioration pour éventuellement arriver à en obtenir en milieu non-contrôlé. Enfin, ce premier projet d'intégration et d'évaluation de capacités interactives d'un robot mobile servira à alimenter une prochaine itération avec un robot plus sophistiqué présentement en conception.

Mots-clés : interaction humain-robot, interruption, robot mobile autonome, multiples modalités d'interaction

REMERCIEMENTS

Je remercie François Michaud pour son soutien et ses conseils pendant toute la durée de ma maîtrise. Mais aussi à tous mes collaborateurs au sein du laboratoire IntRoLab, et plus particulièrement François Ferland et Dominic Létourneau avec qui j'ai travaillé plus étroitement.

Merci également à David Brodeur qui m'a assisté durant les expérimentations avec le robot, ainsi qu'aux volontaires qui ont participé aux essais.

Je remercie le Conseil de recherche en sciences naturelles et en génie (CRSNG) pour son soutien financier, sans quoi ce travail n'aurait pu être effectué dans les mêmes conditions.

Je suis reconnaissant des amitiés qui se sont créées au Québec, ainsi qu'à Camille sans qui je n'aurais pas traversé cette expérience avec la même envie. Je suis également heureux de pouvoir compter sur le soutien sans faille de ma famille et amis malgré l'éloignement.

TABLE DES MATIÈRES

| | | |
|----------|--|-----------|
| 1 | Introduction | 1 |
| 2 | Interaction humain-robot humanoïde | 3 |
| 2.1 | Modalités gestuelles | 6 |
| 2.2 | Modalités visuelles | 6 |
| 2.3 | Modalités audio | 7 |
| 2.4 | Mobilité et décision | 9 |
| 2.5 | Types de recherche | 9 |
| 2.5.1 | Démonstration | 9 |
| 2.5.2 | Évaluation | 11 |
| 3 | Article | 15 |
| 3.1 | Avant-propos | 15 |
| 3.2 | Abstract | 17 |
| 3.3 | Introduction | 17 |
| 3.4 | Johnny-Jr, an Interactive Omnidirectional Robot Platform | 19 |
| 3.4.1 | Robot Control Architecture | 21 |
| 3.4.2 | Interaction Scenario | 24 |
| 3.5 | HRI Evaluation with Johnny-Jr | 27 |
| 3.5.1 | Within-Subject Study in Controlled Conditions | 28 |
| 3.5.2 | Between-Subject Study in Open Settings | 34 |
| 3.6 | Discussion and Future Work | 39 |
| 3.7 | Acknowledgements | 40 |
| 4 | Démarche et contribution | 41 |
| 5 | Conclusion | 45 |
| A | Images des robots existants | 47 |
| B | Demande au comité d'éthique | 57 |
| C | Questionnaire pour les expérimentations en milieu contrôlé | 69 |
| D | Grille d'évaluation et questionnaire pour les expérimentations en milieu non-contrôlé | 71 |
| | LISTE DES RÉFÉRENCES | 75 |

LISTE DES FIGURES

| | | |
|------|--|----|
| 2.1 | Disposition des émotions suivant les données recueillies [12] | 7 |
| 2.2 | Flot d'information entre le module de reconnaissance de l'engagement et l'architecture décisionnelle du robot [52] | 8 |
| 2.3 | Architecture ACT-R/E et sa relation avec l'environnement [70] | 10 |
| 3.1 | Johnny-Jr robot | 20 |
| 3.2 | Johnny-Jr facial expressions | 21 |
| 3.3 | Johnny-Jr gestures | 21 |
| 3.4 | Control architecture of Johnny-Jr | 22 |
| 3.5 | Finite-state machine of the Interaction Scenario module | 24 |
| 3.6 | Trajectory of Johnny-Jr when approaching a person | 27 |
| 3.7 | Q1 preferences for the test population | 31 |
| 3.8 | Q1 preferences for group A (left side - blue) and for group B (right side - yellow) for each configuration | 31 |
| 3.9 | Q2 preferences for the test population | 32 |
| 3.10 | Q2 preferences for group A (left side - blue) and for group B (right side - yellow) for each modality | 33 |
| 3.11 | Johnny-Jr in position to interact with people in open settings, for a configuration not involving mobility (with the robot being connected to its charger) | 35 |
| 5.1 | Illustration de Johnny-0 | 46 |
| A.1 | Armar | 47 |
| A.2 | Asimo | 47 |
| A.3 | Autom | 48 |
| A.4 | Bandit II | 48 |
| A.5 | Barthoc | 48 |
| A.6 | Cramer | 49 |
| A.7 | Fritz | 49 |
| A.8 | George | 49 |
| A.9 | GRACE | 50 |
| A.10 | HRP-2W | 50 |
| A.11 | iCub | 50 |
| A.12 | Justin | 51 |
| A.13 | Kaspar | 51 |
| A.14 | Keepon | 51 |
| A.15 | Kismet | 52 |
| A.16 | Leonardo | 52 |
| A.17 | MDS | 52 |
| A.18 | Mel | 53 |
| A.19 | Melvin | 53 |
| A.20 | Microsoft | 53 |

| | |
|---------------------------|----|
| A.21 PR2 | 54 |
| A.22 Rackham | 54 |
| A.23 Robotinho | 54 |
| A.24 Robovie II | 55 |
| A.25 Robovie IV | 55 |
| A.26 RoboX | 55 |
| A.27 SIG | 56 |
| A.28 Spartacus | 56 |

LISTE DES TABLEAUX

| | | |
|-----|--|----|
| 2.1 | Capacités d'interaction de robots interactifs humanoïdes | 5 |
| 3.1 | Dialogue used for the interaction scenario | 26 |
| 3.2 | States duration of the interaction scenario | 26 |
| 3.3 | Difference ($ \bar{R}_u - \bar{R}_v $) of average ranks of preferences between each configuration (Q1) | 30 |
| 3.4 | Difference ($ \bar{R}_u - \bar{R}_v $) of average ranks of preferences between each modalities (Q2) | 32 |
| 3.5 | Q3 results | 34 |
| 3.6 | Results for Q8 | 37 |
| 3.7 | Data gathered with the 99 participants of the between-subject study in open settings | 38 |

CHAPITRE 1

Introduction

Pour qu'un robot arrive à interagir de manière naturelle avec un humain, il doit pouvoir intégrer différentes modalités d'interaction telles que la parole, le suivi du regard, les gestes, les expressions et la mobilité. En soit, ceci peut sembler simple, car ce sont des habiletés qui sont utilisées de manière intuitive chez l'humain. Et pourtant, pour un robot mobile humanoïde, cela reste un défi de taille, car chacune de ces modalités entraîne des contraintes particulières qui demandent de relever un défi d'intégration de taille. À cette conception s'ajoute le défi d'expérimentation des capacités interactives développées avec des personnes, rattaché au domaine d'étude *Human-Robot Interaction* (HRI), ce qui en soit implique un domaine de recherche qui va au-delà de l'ingénierie. Étant à ces débuts, le domaine du HRI consiste principalement à l'expérimentation avec des prototypes utilisés en milieux contrôlés (laboratoire) avec des études quantitatives, ou en milieux ouverts avec des études qualitatives non structurées. Il en résulte des recherches de types exploratoires qui examinent principalement la complexité technologique des modalités interactives à mettre en œuvre, et non l'impact de ces modalités sur la qualité des interactions. De telles recherches demandent des compétences en évaluation et en études de facteurs humains (e.g., psychologie, sciences humaines) qui débordent des compétences en robotique, ce qui nécessitent des collaborations interdisciplinaires.

Par ce travail, le but est de faire une première itération d'un processus de conception et d'évaluation de modalités interactives avec un robot humanoïde. Plus spécifiquement, l'intérêt réside dans l'intégration de capacités d'interaction vocales avec les expressions faciales, le suivi du regard, les gestes et la mobilité. La mobilité apparaît particulièrement intéressante, car elle a été peu étudiée jusqu'à maintenant. Le robot Johnny-Jr a été spécifiquement conçu pour réaliser cette étude en intégrant différents projets du laboratoire, dont une base omnidirectionnelle à roues orientables [21], un bras interactif [38] et un visage robotique capable d'expressions faciales [42], et en programmant les modalités d'interaction requises avec l'exploitation de capteurs tels qu'un détecteur de proximité laser et d'une caméra Kinect. Du point de vue expérimental, cette démarche innove en effectuant une étude entre participants (*within subject study*) en milieu contrôlé ainsi qu'une étude inter-participants (*in-between subject study*) en milieu non-contrôlé afin d'évaluer les quatre configurations de modalités interactives suivantes :

- Le robot interagit seulement avec la voix et les expressions faciales.
- Le robot utilise la voix, les expressions faciales et l'orientation du regard.
- Le robot utilise la voix, les expressions faciales, l'orientation du regard et les gestes.
- Le robot utilise la voix, les expressions faciales, l'orientation du regard, les gestes et le déplacement.

Dans ces configurations, le robot doit arriver à attirer l'attention d'une personne et d'engager une interaction le menant à aller chercher un objet (une casquette) qui se trouve au sol non loin du robot et la lui remettre. L'hypothèse à valider est que la qualité d'interaction sera meilleure avec l'accroissement du nombre de modalités d'interaction sur le robot. Pour les expérimentations en milieu contrôlé, les participants sont mis au courant de l'intention du robot, tandis qu'en milieu non contrôlé les personnes sont interpellées sans savoir ce que le robot souhaite accomplir. L'objectif de notre démarche est d'arriver à mettre en place une méthodologie pour l'évaluation rigoureuse avec des mesures quantitatives et qualitatives de l'interaction sur le moment.

La contribution scientifique résultantes du présent mémoire est donc double :

1. Démontrer la faisabilité d'intégrer les modalités interactives dans un processus décisionnel pour le robot. Le défi est important ici, car l'orientation du regard et des gestes sont influencés par la mobilité du locuteur et du robot, et demande donc une autonomie décisionnelle plus grande de la part du robot. Cette démonstration est source de pistes utiles pour la conception de robots interactifs autonomes.
2. Expérimenter avec différents protocoles afin de dégager une procédure menant vers des études rigoureuses en milieux naturels.

Le présent travail se présente comme un mémoire par article et est organisé de la façon suivante. Le chapitre 2 présente la revue de l'état de l'art sur les robots humanoïdes interactifs ainsi que leurs évaluations, qui a servi à élaborer le projet de recherche pour le mémoire. Le chapitre 3 présente l'article soumis à l'*International Journal of Social Robotics* sur les travaux d'intégration des modalités du robot et d'évaluation de ses interactions avec les personnes. Tous les éléments du projet y sont couverts. Enfin, le chapitre 4 conclut le mémoire.

CHAPITRE 2

Interaction humain-robot humanoïde

Il est clair qu'il existe une multitude de façons d'interagir entre un humain et un robot dans de multiples contextes. Il est possible de catégoriser ces interactions suivant différentes dimensions.

En psychologie de la communication et de l'interaction interpersonnelle, on décompose les moyens de communication en deux groupes : verbal et non-verbal [18]. Ces deux catégories sont bien distinctes et desservent des objectifs bien différents. En effet, les messages verbaux font passer les idées et les pensées tandis que les messages non-verbaux transmettent les sentiments et les désirs. La caractéristique principale des messages non-verbaux (e.g., gestes, regard, toucher, intonation de la voix) est qu'ils sont tout le temps présents, forment un tout (tout le corps parle d'une même voix) et sont difficilement contrôlables. Dans une conversation, on évalue à 65% l'information passée par des messages non-verbaux. De plus, lorsqu'il y a une contradiction entre les messages verbaux et non-verbaux, ce sont ces derniers qui sont pris en considération, car ils sont plus fiables. L'inconvénient, c'est qu'ils sont moins clairs que les messages verbaux.

Cette distinction entre messages verbaux et non-verbaux mène à proposer une classification par modalités, en lien avec le matériel utilisé pour mettre en œuvre ces modalités. D'autres dimensions possibles sont la contribution scientifique du travail ainsi que l'application visée.

Catégoriser les recherches dans ces trois classes ne s'avère pas simple, car le robot peut être capable d'utiliser une multitude de modalités (dont certaines inusitées) et d'effectuer plusieurs types d'applications. De plus, la contribution scientifique peut être difficile à cibler étant donné la diversité des sujets abordés et le niveau de profondeur des études réalisées.

Ce chapitre présente une catégorisation des travaux de recherche de ce domaine selon les modalités d'interaction et le type de recherche réalisé (e.g., démonstration et évaluation) afin de mettre en évidence l'intégration des modalités d'interaction et la nature de la contribution scientifique réalisée. Une catégorisation des travaux selon le type d'application n'a pas été faite, car le projet ne vise pas un domaine particulier.

Le tableau 2.1 présente la catégorisation réalisée pour les robots illustrés en photo à l'annexe A. Les modalités d'interaction sont regroupées selon les modalités gestuelles, visuelles, audio ainsi que la mobilité et la décision. Dans ces groupes, les modalités sont présentées en ordre suivant leur type de recherche (e.g., démonstration et évaluation) et ensuite suivant le nombre de modalités utilisées, du plus petit au plus grand. La grande majorité des robots interactifs possèdent, sous une forme ou une autre, des modalités gestuelles, soit avec des expressions faciales ou des mouvements corporels. Parmi les modalités gestuelles, c'est celle du toucher qui semble être la moins exploitée. Cette modalité est utilisée pour manipuler des objets comme dans le cas de Justin [50], l'iCub [72] et Leonardo [25]. Il y a aussi Robovie [44] qui utilise 56 capteurs tactiles positionnés sur son corps pour détecter lorsqu'une personne le touche. Une brève analyse des différents aspects considérés par les robots interactifs humanoïdes est présentée dans les prochaines sous-sections.

Tableau 2.1 Capacités d'interaction de robots interactifs humanoïdes

| Robot | Gestes | | | Vision | | | | | | Audio | | | | Mobilité et décision | | | Type de recherche ¹ | Référence |
|-----------------|--------------------|--------------------|---------|------------------------------|--------------------------|---------------------|-------------------------|---------------------|---------------------|-----------------------|-----------------------|-----------------|---------------------|----------------------|-----------------------|----------|--------------------------------|------------------|
| | Expression faciale | Mouvement corporel | Toucher | Détection et suivi de visage | Reconnaissance de gestes | Détection du regard | Reconnaissance d'objets | Écran d'information | Localisation du son | Interprétation du son | Reconnaissance vocale | Synthèse vocale | Gestion de dialogue | Emotion vocale | Perception avec laser | Mobilité | Architecture décisionnelle | |
| Asimo (Nakadai) | | X | | | | | | | X | | X | | | | | - | Démonstration | [47] |
| Kismet | X | X | | X | | | X | | | | | | | | | - | Démonstration | [10, 12] |
| ARMAR | | | | | | | X | | | | X | | | | | Roues | Démonstration | [2, 3] |
| SIG | | | | X | | | | | X | | X | X | | | | - | Démonstration | [48, 49] |
| Justin | | | X | | | | X | | | | X | | | | | Roues | Démonstration | [50] |
| Bandit II | X | X | | | X | | | | | | | X | | | | Roues | Démonstration* | [68] |
| PR2 | X | X | X | X | | | | | | | | | | | X | Roues | Démonstration* | [67] |
| GRACE | X | | | | | | | X | | | X | X | | | X | Roues | Démonstration | [39, 64] |
| Leonardo | X | X | X | X | | X | X | | | | X | | | | | - | Démonstration | [25] |
| RoboX | X | | | X | | | | | | | X | X | | | X | Roues | Démonstration | [63] |
| George | X | | | X | | | | | X | | X | X | | | X | Roues | Démonstration | [71] |
| Fritz | X | X | | X | | | | | X | | X | X | | X | | - | Démonstration | [6] |
| iCub | X | X | X | X | | | X | | | | X | X | | | | - | Démonstration* | [72] |
| Octavia (MDS) | X | X | | X | X | X | X | | | | | | | | | - | Démonstration* | [70] |
| Barthoc | X | X | | X | X | | | | X | | X | X | | | | - | Démonstration | [65] |
| Melvin (Rich) | X | X | | X | X | X | X | | | | | X | X | | | - | Démonstration* | [52] |
| Spartacus | | | | | | | X | X | X | | X | X | | | X | Roues | Démonstration | [41] |
| Rackham | X | | | X | | | X | X | | | X | X | | | X | Roues | Démonstration | [15] |
| HRP-2W | | | X | | X | | X | X | X | X | | | | | X | Roues | Démonstration | [28] |
| Microsoft | X | | | X | X | X | | X | X | | X | X | X | | | - | Démonstration* | [9] |
| Robotinho | X | X | | X | X | X | X | | | | X | X | X | | X | Roues | Démonstration | [20] |
| Kaspar | X | X | | | | | | | | X | | | | | | - | Évaluation* | [32, 33] |
| Keepon | X | X | | X | | | X | | | X | | | | | | - | Évaluation | [34] |
| Auton | | X | | X | | | | X | | | X | X | | | | - | Évaluation | [31] |
| Mel | | X | | X | X | X | | | X | | X | X | X | | | - | Évaluation | [58-61] |
| Robovie | | X | X | X | | | | | X | | X | X | X | | X | Roues | Évaluation | [29, 44, 45, 78] |
| Total (26) | 15 | 16 | 6 | 18 | 8 | 6 | 10 | 6 | 10 | 3 | 17 | 16 | 5 | 1 | 8 | 12 | 8 | |

¹ Les travaux annotés d'une étoile sont en cours.

2.1 Modalités gestuelles

En psychologie, les gestes sont classés en quatre catégories [18] :

- Emblématique, c'est à dire, propre à une communauté et que tout le monde connaît. Le robot Fritz [6] est capable d'effectuer plusieurs gestes emblématiques (e.g., saluer, demander une question et demander de venir plus près).
- Illustratif, c'est à dire, qui accompagne la parole. Le robot Mel [61] est capable de faire des gestes illustratifs comme pointer du doigt ou du regard et acquiescer.
- Régulateur, c'est à dire, qui coordonne les échanges. Ce type de geste est complexe à mettre en œuvre sur un robot, car ces gestes sont instinctifs pour l'humain (e.g., toucher son interlocuteur pour prendre la parole). Les travaux de Trafton [71] avec George témoignent des premières avancées pour qu'un robot puisse suivre une discussion.
- Émotif, c'est à dire, qui exprime les émotions (principalement sur le visage). Cette émotion est véhiculée par les expressions faciales qui sont de plus en plus utilisées sur les robots récents. Le robot Kismet [12] du MIT a été un des premiers robots à les utiliser en exprimant ses émotions selon l'excitation, la position et l'intérêt de l'interlocuteur, générés à partir de détection de visage, de couleur et de mouvement tel que décrit à la figure 2.1. Les émotions peuvent aussi s'exprimer par un visage affiché sur un écran [9, 40, 64, 71]. L'utilisation d'un visage virtuel n'est pas à privilégier dans le cadre d'une interaction humain-robot, car l'écran tend à être moins engageant à la conversation qu'une tête robotisée [76].

2.2 Modalités visuelles

Les modalités visuelles sont utilisées pour donner au robot des informations sur son environnement visuel. Ces modalités sont largement utilisées pour pouvoir agir de façon cohérente aux actions de l'interlocuteur. La détection et le suivi de visages sont les modalités visuelles les plus fréquemment utilisées. Les travaux de Viola et Jones [75] sur la détection de visages ont été intégré à OpenCV. Cela facilite leurs utilisations sur les robots tels que Fritz [6], Barthoc [65] et Rackham [15]. La reconnaissance d'objets est aussi utilisée sur les robots interactifs. Cette reconnaissance se fait principalement à partir de la couleur [12, 25, 52]. Certains robots utilisent des algorithmes plus complexes comme SIFT [3] pour que le robot soit capable de repérer et ensuite utiliser des objets de la vie courante.

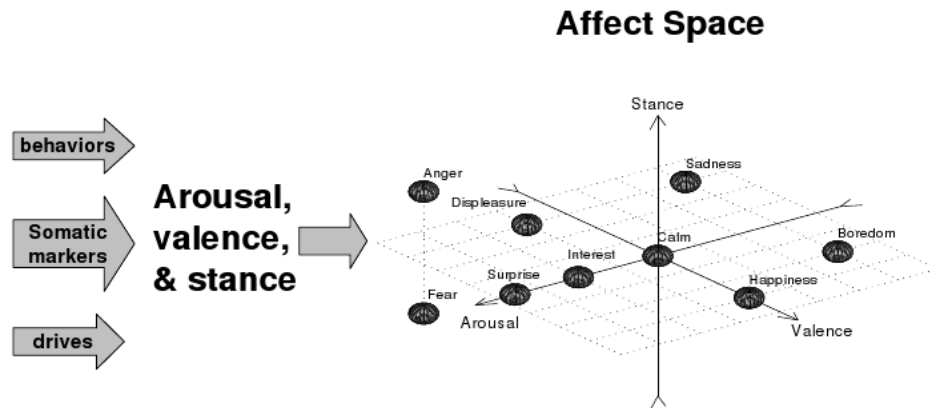


Figure 2.1 Disposition des émotions suivant les données recueillies [12]

D'autres modalités ont commencé à être utilisées sur les robots présentés dans le tableau 2.1. En effet, quelques robots sont capables de détecter les gestes ou la direction du regard. La détection de la direction du regard donne au robot la capacité de connaître ce qui intéresse son interlocuteur à chaque instant. Mais cette modalité est encore peu présente due à sa complexité technique. Le robot réceptionniste [8] de Dan Bohus est capable de savoir vers qui est orientée l'attention d'une personne (e.g., le robot, une certaine personne). Le robot Mel est par exemple capable de détecter les mouvements de la tête et ensuite de reconnaître s'il y a un acquiescement en utilisant les informations du contexte [46]. De même, un modèle capable de détecter l'engagement entre un robot et une personne a été développé récemment sur le robot Melvin [52]. Ce modèle se base sur les données recueillies par les différentes modalités (les gestes, les regards, la parole autant pour l'humain que pour le robot), comme le montre la figure 2.2. Ceci est une démarche originale visant la détection de l'engagement sans être dépendant des modalités d'interaction.

Sur certains robots interactifs, des écrans sont aussi utilisés pour afficher ou demander des informations à la personne avec qui le robot interagit [31, 41].

2.3 Modalités audio

La reconnaissance vocale et la synthèse vocale sont les modalités audio qui se retrouvent le plus fréquemment sur les robots interactifs, probablement, car elles sont à la base des interactions naturelles avec l'humain. Plusieurs utilisent des systèmes de reconnaissance

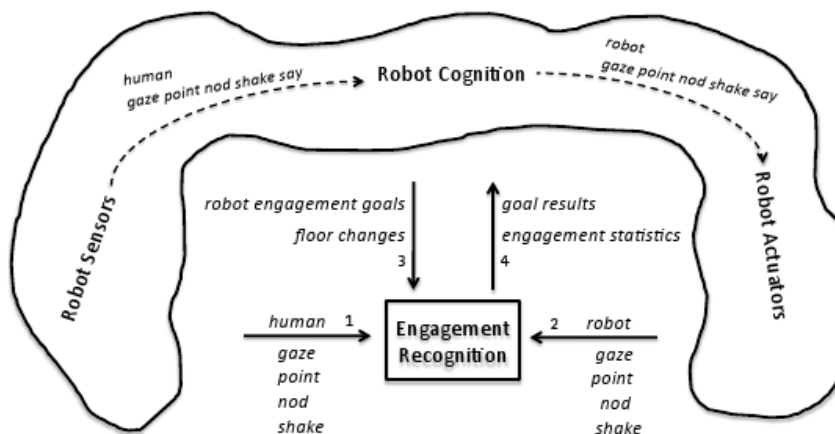


Figure 2.2 Flot d'information entre le module de reconnaissance de l'engagement et l'architecture décisionnelle du robot [52]

comme Sphinx [77] et Nuance¹, et des systèmes de synthèse vocale comme Festival [69] et MBROLA [19].

La capacité de localiser les sons autour du robot se retrouve aussi de plus en plus sur les robots interactifs, avec des systèmes issus du milieu de la recherche (e.g., ManyEars [74]) et des systèmes commerciaux (e.g., le Voice Tracker de Acoustic Magic²). L'avantage de ManyEars est qu'il permet la localisation sur 360 degrés alentours du robot, et que cette localisation sert à séparer les sources sonores pour une meilleure reconnaissance vocale.

La plupart des robots interactifs qui peuvent dialoguer avec des interlocuteurs humains le font à partir de scripts pré-établis ou encore des gestionnaires de dialogues, tels que ceux utilisés sur Cramer avec ADE [13] ou sur le robot de Microsoft [8]. Un gestionnaire de dialogue permet de définir des modèles d'échanges généraux adaptés à un contexte particulier comme la présentation d'un objet avec le robot Mel [60] en utilisant Collagen [53] pour gérer le dialogue. CSLU est le gestionnaire de dialogue utilisé sur Spartacus [43].

Enfin, un petit nombre de robots sont en mesure d'interpréter le son, pour détecter des bruits, comme HRP-2W [28] qui reconnaît des objets grâce à leur son caractéristique), ou d'identifier des émotions dans la voix³.

¹www.nuance.com

²<http://www.acousticmagic.com/>

³Un collaborateur d'IntRoLab, Pierre Dumouchel de l'École de technologie supérieure, travaille sur cette capacité.

2.4 Mobilité et décision

La grande majorité des robots interactifs sont statiques, c'est à dire qu'ils sont fixés sur une base non mobile. La mobilité amène la difficulté supplémentaire de positionner le robot par rapport à l'interlocuteur et son environnement, d'où l'utilisation par certains de capteurs lasers pour la navigation. Les roues sont le mode de locomotion principalement utilisé pour les robots interactifs, et servent à approcher des personnes pour communiquer [55]. Les robots peuvent aussi servir de guide dans une exposition ou un musée, comme c'est le cas pour Rackham [15] et RoboX [63].

Cette mobilité est souvent amenée par une capacité décisionnelle accrue du robot. Cette capacité peut être modélisée selon le processus cognitif humain, tel que ACT-R [70, 71] avec son implémentation réalisée par Trafton et illustrée à la figure 2.3. Elle peut aussi être basée sur des principes de contrôle intelligent, telle que l'architecture à trois niveaux utilisés par Robovie IV [44] : le premier niveau récupère les informations des capteurs et contrôle les actionneurs ; le second gère les événements et sert d'interface entre le premier et le troisième niveau ; le troisième niveau est composé de processus inter-reliés pour le dialogue et le suivi de visage. Certaines architectures décisionnelles sont basés sur les fonctionnalités à intégrer [2, 64]. D'autres sont des extensions basées sur l'approche comportementale [4].

2.5 Types de recherche

Le type de recherche fait référence à l'objectif de recherche poursuivi et la méthodologie expérimentale qui en découle. Tous les travaux présentés au tableau 2.1 portent sur des modalités interactives validées sur des robots, soient pour démontrer la faisabilité d'exploiter ces modalités (seules ou intégrées), soient pour évaluer l'impact (qualité, efficacité, utilisabilité, etc.) de ces modalités lors de l'interaction humain-robot.

2.5.1 Démonstration

Les recherches de type démonstration sont généralement des preuves de concept qui ont pour objectif de montrer l'efficacité d'une nouvelle modalité ou d'intégrer beaucoup de modalités ensemble sans en faire une étude approfondie ou à long terme. Par exemple, le robot Fritz [6] intègre des modalités audio, visuelles et gestuelles dans le but de faire une présentation. Ce dernier robot a été utilisé lors de plusieurs événements ou expositions, mais aucune mesure concrète n'a été effectuée autant pour mesurer l'efficacité des différentes modalités ou la qualité de l'interaction. Seulement l'impression de satisfaction

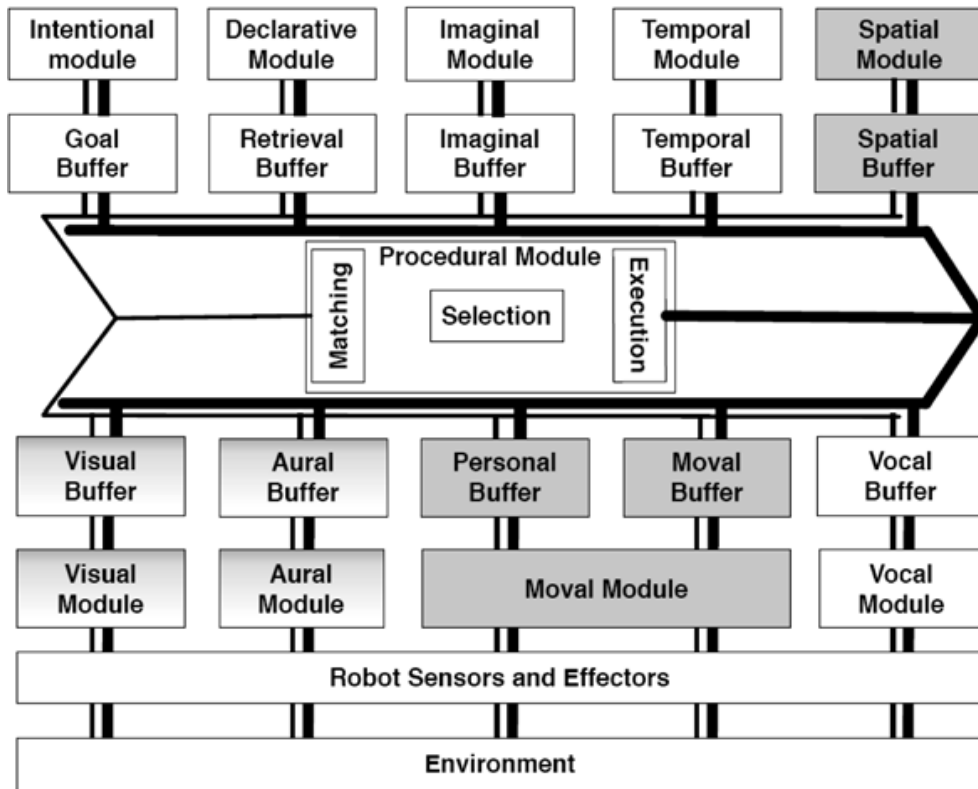


Figure 2.3 Architecture ACT-R/E et sa relation avec l'environnement [70]

générale a été recueillie. Il faut retenir que des expositions sont bien pour tester un robot, mais il est difficile d'y conduire des évaluations.

Les travaux de Nakadai et al. [47] sur le robot Asimo ou de Okuno et al. [49] sur le robot SIG démontrent quant à eux l'efficacité de la modalité permettant la localisation et l'extraction de différentes sources sonores. Dans les deux cas, un scénario précis et répétable a été expérimenté. Dans un cas, le robot Asimo est considéré comme un juge pour un jeu de roche-papier-ciseaux. Il doit décider du vainqueur entre les deux ou trois participants et le vainqueur n'est pas décidé tant qu'il y a égalité. Le robot a été évalué suivant deux critères : le taux de succès pour un jugement et le taux d'accomplissement de la tâche. De plus, la position des participants entre eux change de 30 degrés tous les dix jeux. Il a ainsi été joué 330 jeux de roche-papier-ciseaux. Le résultat est un taux bon jugement de 75% et d'accomplissement de la tâche de 70%.

Bien que certains travaux démontrent une volonté d'évaluer la qualité des interactions, ils restent au stade de la démonstration. Par exemple, la méthodologie abordée dans les travaux sur le robot Melvin [52] montre une volonté de reproduire le genre d'interaction propre au comportement humain. En effet, le modèle qui permet de reconnaître l'engagement sur le robot Melvin a été précédé par des heures d'analyse vidéo sur des scènes

mettant en interaction deux personnes. Il en a été déduit qu'il y a quatre types d'évènements suscitant l'engagement : le regard directeur, le regard mutuel, la conversation adjacente en pair et la notification de retour. Cela a été inclus dans le modèle et un scénario nommé *The Pointing Game* a été mis en place pour tester le modèle. Les évaluations sont en cours pour connaître la qualité de ce modèle. Pour l'instant toutefois, les travaux présentés montrent uniquement le fonctionnement et l'intérêt d'un module indépendant pour la détection de l'engagement entre l'humain et le robot. Pour le robot MDS, Trafton et al. [70] se basent sur une expérimentation de 1998 sur le suivi du regard par 63 enfants de 6 à 11 mois. Lors de leurs travaux, ils reproduisent l'expérimentation avec le modèle cognitif de suivi du regard du robot afin de pouvoir comparer les résultats avec ceux de l'étude de 1998. Le critère qui est pris en compte dans cette étude est le taux de réussite à suivre le regard. Les travaux effectués rendent la modalité du suivi du regard proche du comportement humain.

2.5.2 Évaluation

Une fois les modalités intégrées au robot suivant l'application désirée, la suite logique consiste à mettre en place un scénario qui permet d'évaluer la qualité de l'interaction. On peut voir au tableau 2.1 qu'il y a très peu de recherche de type évaluation. Cela peut en partie s'expliquer par le fait que beaucoup de projets sont en cours et peu d'évaluations ont été présentées, ou encore par la complexité inhérente de réaliser l'intégration des modalités interactives.

Il existe un article de référence qui définit les métriques à utiliser pour évaluer une interaction humain-robot (HRI) [66]. Cet article présente des classes de métriques regroupées suivant les tâches possibles en HRI afin que soit adoptées une taxonomie pouvant mener à des comparaisons entre les travaux. Par exemple, des métriques générales sont définies pour évaluer l'interaction humain-robot. Les deux métriques principales pour cela sont la performance du système et son évaluation subjective. Celles-ci sont reprises dans les travaux de Sidner et al. [59] et Kose-Bagci et al. [32].

Les travaux de Takayama sont fondés pour leur part sur la recherche à base de variables pour évaluer une HRI [67]. Par exemple, dans [22], Takayama examine l'influence de la forme (une voiture versus un robot humanoïde) et de l'assemblage (assemblée en lego par l'utilisateur ou par une tierce partie) en HRI. Des questions sont ensuite posées aux personnes pour connaître l'impact des variables sur leur ressenti. Par exemple, une question est «Que ressentiriez-vous si le robot était détruit ?» et la personne peut répondre sur une

échelle de 1 à 5, de horrible à plaisant. Cette question permet de déterminer l'attachement de la personne au robot.

Sidner et al. [59] ont mis en place une expérimentation pour évaluer de façon rigoureuse l'engagement entre un humain et un robot. Le scénario consiste à la présentation d'un objet par Mel à une personne. Dans ce scénario, Mel a été évalué avec deux comportements. Pour le premier comportement, le robot interagit avec la personne seulement avec la parole. Pour le deuxième, Mel communique avec la parole et les gestes. Il y a eu 27 personnes qui ont participé à l'expérience d'un des deux comportements.

Pour évaluer les dialogues entre humain et robot, les personnes ont répondu à des questions sur la qualité de l'interaction. Les personnes semblent en général satisfaites de la qualité de l'interaction et la compare à un humain. Des différences ont été observées sur la réaction des personnes entre les deux comportements de Mel. Lorsque Mel utilise les gestes, l'interlocuteur est plus impliqué dans la conversation. Par exemple, lorsque Mel indique de regarder la table, 37.5% des personnes vont la regarder. Ce pourcentage monte aussi à 63% lorsque Mel tourne aussi la tête pour regarder la table.

Une expérience a été menée avec Robovie II pour déterminer l'importance de la direction du regard [27]. Pour cela, Robovie devait montrer une affiche à une personne. Deux séries de dix évaluations ont été faites. Lors de la première, Robovie montre une affiche en la pointant et la regardant. Lors de la deuxième, Robovie pointe seulement l'affiche. L'expérience a démontré que la direction du regard est très importante. Une seconde expérience, à long terme cette fois, fut aussi menée avec Robovie IV [44]. Cette expérimentation consiste à faire fonctionner le robot sur 150 mètres dans le laboratoire où circulent plus de 100 personnes pendant 6 mois. Ensuite, une évaluation statistique sur une durée de six semaines a été faite [45]. Il y a eu 23 participants et 76 questionnaires remplis. Cette expérimentation a pour but de déterminer les facteurs d'acceptation du robot dans l'environnement humain. Les analyses montrent que la génération d'un comportement approprié par le robot est le plus important pour qu'il soit accepté par la communauté. La familiarité et le rôle du robot sont aussi à prendre en compte.

Quelques autres expériences à long terme ont aussi été réalisées avec Keepon [34] et Autom [31]. Keepon [34] est un robot créé pour interagir avec les enfants de façon simple, naturelle et non verbale. Il est destiné à l'étude des enfants et de leurs comportements. Des expériences avec Keepon ont été faites avec plusieurs dizaines d'enfants, âgés de quelques mois à cinq ans, autant en laboratoire que dans les écoles. Dans une école, Keepon a été introduit par session de 90 minutes dans une classe de 27 enfants et cela pendant 20 sessions,

afin d'identifier les différentes sortes de comportement de l'enfant vis-à-vis de Keepon : violent ou protecteur, donneur, pointeur et attribuant une intentionnalité propre au robot. Les résultats montrent qu'au bout de quelques mois avec Keepon, un enfant autiste est de plus en plus familier avec le robot, ce qui l'amène à partager son état émotionnel. Autom [31], quant à lui, fut conçu pour interagir socialement et à long terme avec l'humain. Son rôle est d'assister les personnes qui l'utilisent dans un programme de perte de poids. Une étude sur six semaines a été faite avec la participation de 45 personnes âgées de 18 à 72 ans. Les participants ont été séparés en trois groupes. Le premier groupe a utilisé le robot Autom, le second a seulement utilisé le programme informatique d'Autom, et le troisième utilise des fiches papier de suivi pour la perte de poids. L'hypothèse émise est qu'Autom amène une source de motivation supplémentaire pour le suivi du programme. Ceci a été vérifiée, car les personnes ont utilisé Autom 50,6 jours en moyenne alors que le programme et l'usage de fiches papier ont été respectivement utilisés 36,2 jours et 26,7 jours.

CHAPITRE 3

Pardon de vous interrompre, mais puis-je avoir votre attention ?

Une étude de cas sur l'influence des modalités verbales et non verbales quand le robot engage l'interaction avec une personne

3.1 Avant-propos

Auteurs et affiliation :

V. Rousseau : étudiant à la maîtrise, Université de Sherbrooke, Faculté de génie, Département de génie électrique et de génie informatique.

F. Ferland : étudiant au doctorat, Université de Sherbrooke, Faculté de génie, Département de génie électrique et de génie informatique.

F. Michaud : professeur, Université de Sherbrooke, Faculté de génie, Département de génie électrique et de génie informatique.

Date de soumission : 23 septembre 2011

Revue : International Journal of Social Robotics

Titre en anglais :

Sorry to Interrupt, But Can I Have Your Attention? A Case Study on the Influence of Verbal and Non-Verbal Modalities When A Robot Engages Interaction With A Human

Contribution au document :

Le chapitre 2 met en évidence l'état de l'art lorsque fut élaboré le projet. À partir des informations relevées, un scénario expérimental a été élaboré en fonction des fonctionnalités disponibles au laboratoire, ce qui mena à la conception du robot Johnny-Jr et des expérimentations décrites dans le chapitre 3 sous la forme d'un article. Cet article contribue au mémoire en décrivant le robot, le scénario, l'évaluation et les résultats du travail réalisé dans le cadre de cette maîtrise. L'annexe B, l'annexe C et l'annexe D présentent les

documents en appui (demande au comité d'éthique, questionnaires et grille d'évaluation) pour les expérimentations réalisées en milieux contrôlé et non-contrôlé.

Résumé français :

Les études d'interaction humain-robot sont complexes, car elles impliquent généralement l'intégration de modalités verbale et non-verbale qui n'ont pas forcément la même influence dépendamment du contexte d'interaction et des personnes avec qui le robot interagit. La conception d'une plateforme humanoïde autonome et interactive demande d'effectuer ce genre d'intégration tout en satisfaisant des contraintes de temps réel et de capacité de calculs. Une telle conception peut bénéficier de la connaissance des influences des modalités à intégrer, ce qui requiert de les tester avec des personnes en conditions réelles. Cet article rapporte une étude de cas conduite avec Johnny-Jr, un robot humanoïde qui possède une base mobile omnidirectionnelle équipée d'un torse et d'un bras avec trois degrés de liberté. Cette étude consiste à observer les influences de la voix, des gestes (visage, tête et bras) et de la mobilité sur l'engagement des personnes pour interagir avec le robot, en conditions contrôlées et non-contrôlées. Les observations faites par ces cas d'études aident à identifier des solutions importantes à mettre en œuvre pour la conception de robots interactifs avancés.

3.2 Abstract

Human-robot interaction studies are complex because they usually involve the integration of verbal and non-verbal modalities which do not necessarily have the same influence depending on the interaction context and the people the robot interacts with. The design of an interactive autonomous humanoid platform requires performing such integration while satisfying real-time and computational constraints. Such design can benefit from knowing the influences of the modalities to integrate, which requires testing them out with people. This paper reports a case study conducted with Johnny-Jr, an omnidirectional mobile base equipped with a humanoid torso and a three degree-of-freedom arm. The study consists of observing the influences of verbal, gestures (facial, head and arm) and mobility in engaging people to interact with the robot, both in controlled conditions and in open settings. Observations made from this case study help identify important issues to tackle for the design of advanced interactive robots.

3.3 Introduction

Our research laboratory is currently engaged in a five-year research initiative involving the complete design and evaluation of an autonomous platform named Johnny-0 [42]. Johnny-0 is a humanoid mobile platform capable of bidirectional interaction in natural ways (through motion, language, touch, affect) with humans, integrating capabilities such as advanced motion, interaction and cognition, and evaluating such capabilities in unconstrained real world conditions. Johnny-0's distinct contributions are in simultaneously exploiting compliance for locomotion and manipulation, in grounding reasoning and actions through behavior-producing modules, and in considering all possible factors experimenting in the wildness of the real world [42].

One situation that we expect Johnny-0 to experience is the necessity to engage someone, whether it is for assistance (e.g., for recharging, grabbing an object out of its reach, get some information) or to fulfill a specific task (e.g., for handing objects, direct people to do specific actions). This involves verbal and non-verbal modalities (eye gaze, gesture, motion) tightly coupled with the physical, perceptual, computational, energetic and acting capabilities of the robot, and allowing the robot to adapt to the unpredictability of interacting with humans.

Following an iterative and incremental design process, we first implemented proof-of-concept prototypes of the robot's subsystems, i.e., omnidirectional and compliant base,

interactive arm using differential elastic actuators, humanoid torso and a control architecture [42]. These prototypes were integrated together to create Johnny-Jr, allowing us to conduct case studies (an important design phase as indicated in [73]) that could provide useful insights influencing the design of Johnny-0 and its use in more in-depth experimentations. Considering that Johnny-0 will have to integrate a high number of sensing, reasoning, processing and acting capabilities on the same platform for interacting with humans in open settings, we figure that conducting a case study on the influence of verbal and non-verbal modalities in human-robot interaction (HRI) would provide key insights on the expected level of performances for the robot's capabilities and on how to conduct evaluation of such capabilities in natural settings.

Our objective with this case study is to examine how a mobile robot can initiate interaction with a human in open settings. This idea was inspired from observing a group of kids selling calendars in a mall during a fundraising activity : kids that were more expressive and moved toward people were more successful in getting people's attention and interest. Therefore, we hypothesize that a robot can be more successful in engaging people to interact when using all of its interaction modalities, especially motion of its mobile base.

Most of the work about communication and attention in human-robot interaction involves the influence of combining verbal communication with non-verbal cues, such as head and gaze control [14, 26, 27, 37], facial features [11, 14, 20], deictic gesture [20, 26, 59] and physical expression [11]. Results suggest that non-verbal cues are beneficial in HRI. These studies usually involve one or two non-verbal modalities using an immobile platform operating in controlled conditions, except for [14] which was conducted in open settings. The work with Robotinho [20] uses a mobile base in public space, but is limited to a demonstration of its capabilities. With mobile platforms, studies have been carried out to : characterize people's personal spaces in regard to a robot [67] ; the approach and positioning of a robot relative to a seated human subject [17] ; or the importance of anticipating human trajectory when a robot approaches people in a mall [30]. The only study we found that involves initiating human-robot interaction is done by Saulnier et al. [56]. Using a Wizard of Oz approach, they examine how a robot can interrupt a human in a socially acceptable manner, observing if humans can interpret the motion and intent of a robot (through motion, speed and head gestures). The participants were asked to rank 10 configurations (i.e., differences in distance of approach, direct or erratic gaze, body rotation and motion speed). Results suggest that interruptibility is mostly influenced by the distance of approach and the robot's speed at close distance, with no significant influence of

gaze and body rotation. They also notice the difficulty of executing the robot's behaviors consistently during experimentation, a limitation when using a Wizard of Oz approach.

In this paper, we present the case study conducted with Johnny-Jr regarding the influences of verbal and non-verbal communication modalities in making a robot autonomously initiate interaction with people, both in controlled conditions and in open settings. The interaction scenario consists of having the robot ask people to give it a baseball cap lying next to it on the floor. Going through the entire design and evaluation process, the objective of this study is to provide data and observations that can influence the design of interactive robots, such as Johnny-0, and their experimentation in natural and open settings.

3.4 Johnny-Jr, an Interactive Omnidirectional Robot Platform

As shown in Fig. 3.1, Johnny-Jr platform integrates an omnidirectional mobile base with an arm installed on a upper torso frame. The mobile base consists of an omnidirectional non-holonomic platform with eight actuators for locomotion [21] : four for propulsion and four for steering the wheels, with each wheel being able to rotate 180° around its steering axis. Propulsion is done using standard DC brushless motors equipped with optical encoders, and capable of reaching 1.47 m/s. Velocity of the mobile base is limited to a maximum of 0.6 m/s for safe and secure approach of people. This maximum was set because in [17], 0.4 m/s was either found to be too slow or just about right, but never too fast. The platform uses steerable wheels motorized using differential elastic actuators (DEA) [35, 36], providing compliance, safety and torque control capabilities. As an underlying safety feature, the mobile base is programmed to stop if an obstacle comes too close to the robot's path. Its hardware architecture consists of distributed modules for sensing and low-level control, communicating with each other through a 1 Mbps CAN bus. A Mini-ITX computer equipped with a 2.0 GHz Core 2 duo processor running Linux is used on-board for high-level control modules. A Hokuyo UTM-30LX laser range finder (30 m range, 270° angle and 25 msec/scan) is placed on the base. Nickel-metal hydrid batteries provide power to the platform for up to 3 hours of autonomy. A wireless gamepad can be used to control the mobile base, to initiate the interaction scenario and to reinitialize odometry.

The humanoid torso comes with two loudspeakers, one Kinect motion sensor (providing a 3D point cloud with color for 640×480 image with an angular field of view of 57°

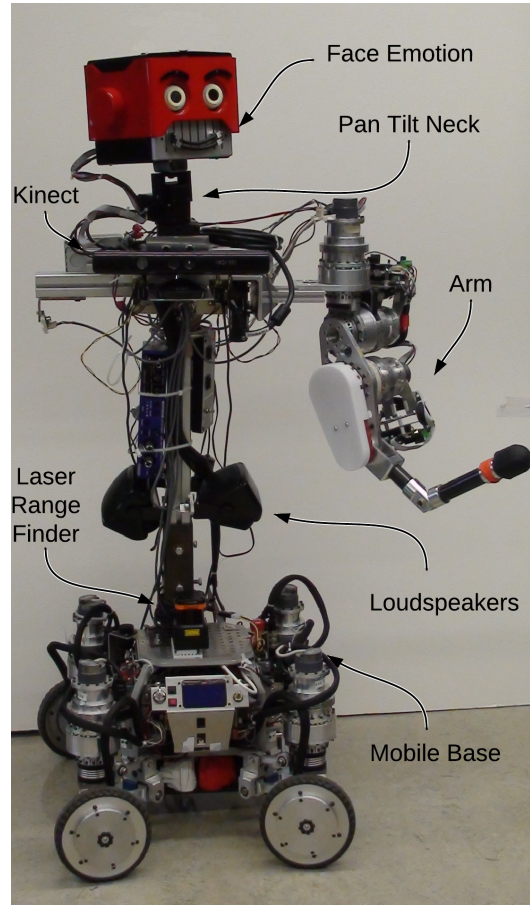


Figure 3.1 Johnny-Jr robot

horizontally and 43° vertically, up to 5 m and at 30 frames per second) and a robot head installed on a pan-tilt unit (PTU-D46-17 of Directed Perception, actuated using step motors : pan $\pm 128^\circ$, tilt $[-17, +25]^\circ$). The robot speaks French using text-to-speech software (l1iaphon [5] for text-to-phoneme conversion, and mbrola [19] for phoneme-to-sound conversion). Figure 3.2 illustrates the facial expressions that the robot can make using nine servo-motors (four for the mouth, three for the eyes and two for the eyebrows). Each emotional state can be expressed by two configurations, interchanging every five seconds to make facial expressions more dynamic.

The arm is attached to the torso and has three serial degree-of-freedom (two in the shoulder and one in the elbow), also actuated with DEA. The arm can sustain impacts with humans or objects, and can be controlled from low to high admittance with gravity compensation [38]. The arm controller receives the desired angles for each joint, and generates spline trajectories between waypoints while ensuring that the arm does not collide with the robot. Gesture programmed on the robot consists of the following, as shown in Fig. 3.3 : Greeting/Goodbye, Come Closer, and Pointing.

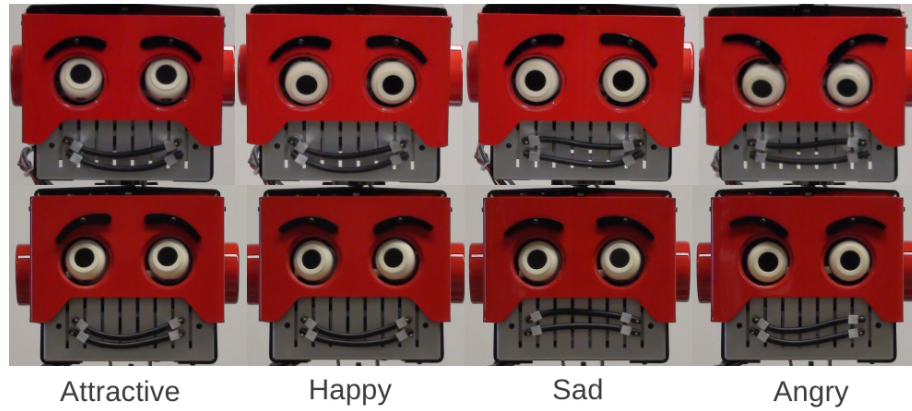


Figure 3.2 Johnny-Jr facial expressions

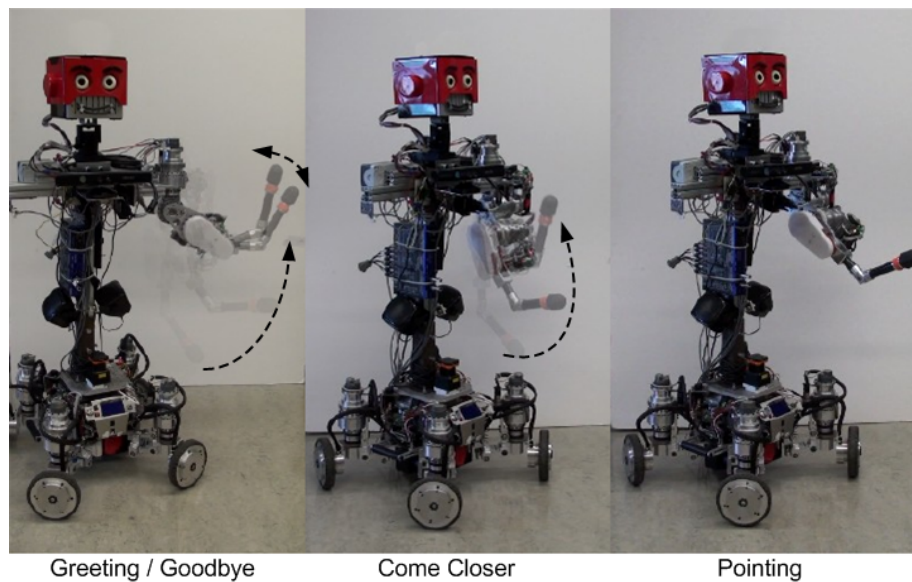


Figure 3.3 Johnny-Jr gestures

3.4.1 Robot Control Architecture

Fig. 3.4 illustrates Johnny-Jr's control architecture implemented using ROS [51]. It consists of a set of behavior-producing (also referred to as behaviors) modules with priority-based arbitration. The **Interaction Scenario** module is a task-level executive system (implemented using SMACH, a Python library integrated in ROS to build state machines [7]) that activates behaviors or inhibits data streams. All modules of the control architecture can access the odometry and joints states of the robot using ROS' *tf* package, which manages and keeps track of every coordinate frames of the robot in the world's coordinate frame. Using data from the Kinect sensor and the OpenNI library integrated in ROS, the **Body Shape Detection** module extracts the centre of mass of people and only tracks the first person detected until it leaves the sensor's field of view. The Kinect sensor's field

of view is however a bit limited when making a robot physically intercept people walking nearby. To provide a wider range of detection, the **Leg Detection** module analyses laser range finder data to identify potential leg positions (with a reliability estimation derived from perceived leg velocity) over 180° and a 5 m range using the approach presented in [1]. Our implementation can take up to 1 sec to validate the presence of a person using leg detection. Position and direction of the first pair of legs detected is provided, and are tracked by the module while they are detected. Combining leg detection with body shape detection allows the robot to first identify potential targets over a large field of view using the laser range finder, to then concentrate on the Kinect which is more reliable for people tracking.

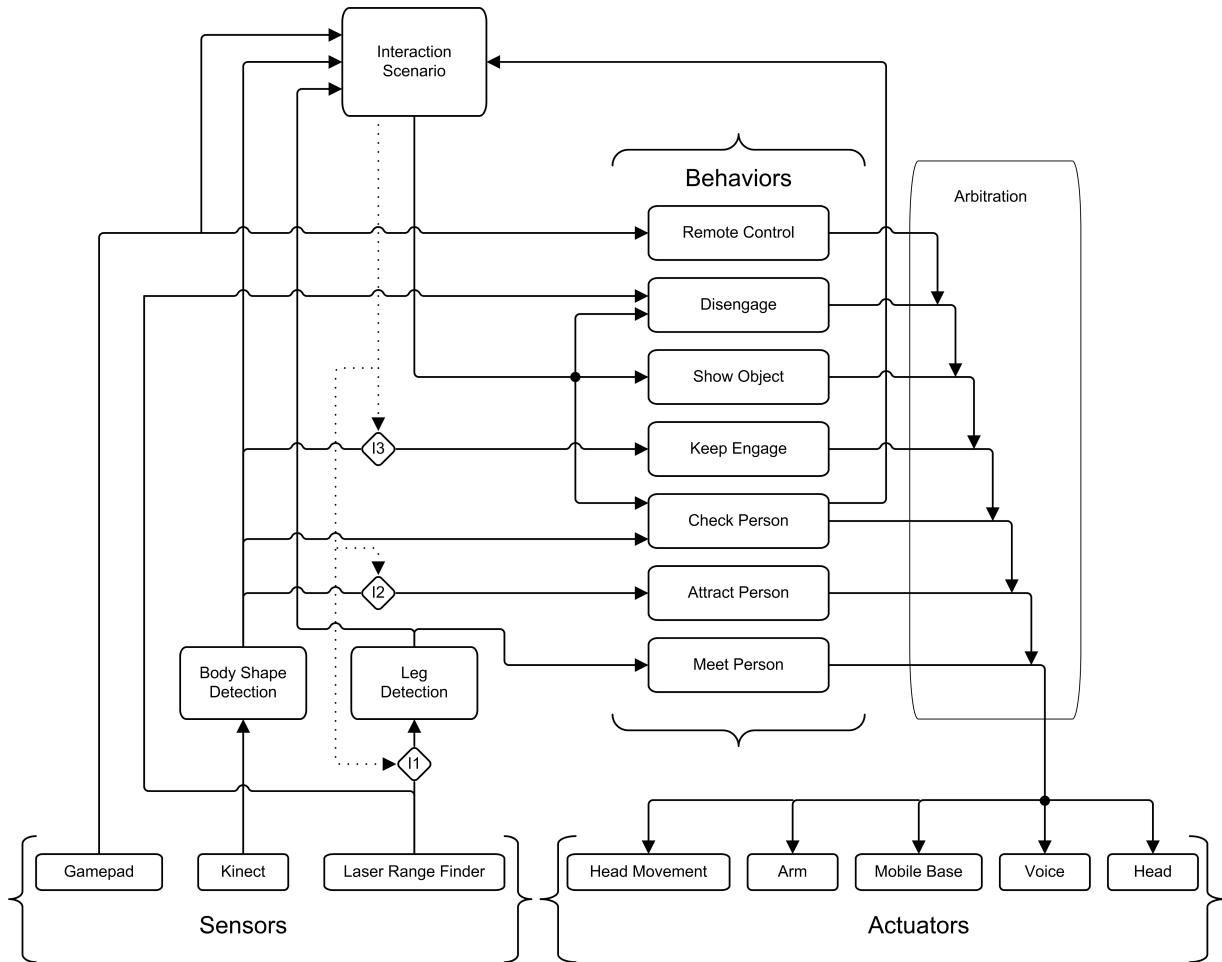


Figure 3.4 Control architecture of Johnny-Jr

Seven behaviors are used to control the robot's actions :

- **Meet Person.** When the legs of a person is detected, this behavior initiates the Greeting gesture with the associated voice message, and makes the robot move to-

ward the person. If the legs are detected on either side of the robot, the behavior generates a curved trajectory (4 m radius trajectory at 0.6 m/s for 3 seconds) with the torso oriented perpendicular to it. No trajectory is generated if the legs are located over $[-45, +45]^\circ$ in front of the robot.

- **Attract Person.** When a body shape is perceived, this behavior allows the robot to maintain a socially acceptable distance (from 1 to 2 m) and keeps the robot facing the person. It can also generate vocal interactions and gestures.
- **Check Person.** This behavior verifies if the person executes the robot's request by positioning it toward the location of the object involved in the interaction scenario (e.g., the baseball cap in our case study).
- **Keep Engage.** Using body shape positions, this behavior tracks the person by controlling the mobile base and the pan-tilt unit. If the person is located $[-5, +5]^\circ$ in front of the robot, only the pan-tilt unit is used, positioning it to follow the centre of the face estimated 1 m above the centre of mass of the person. Otherwise, the behavior makes the robot rotate toward the person's position to assist tracking done by the pan-tilt unit.
- **Show Object.** This behavior makes the robot point toward the position of the object used in the interaction scenario. It does so by generating voice commands, executing the Pointing gesture, orienting the mobile base (rotated by an angle bisector to the angle between the person and the object) and the head toward the object.
- **Disengage.** While generating vocal messages and making the Goodbye gesture, this behavior makes the robot base return to the position it was when engaging interaction. The trajectory is generated using ROS local navigation planner.
- **Remote Control.** Using the gamepad, this behavior allows the experimenter to reposition the mobile base or to stop the robot for security purpose only if necessary.

The activation of **Disengage**, **Show Object** and **Check Person** behaviors by the **Interaction Scenario** is made by specifying a specific goal to achieve. For example, **Check Person** gets activated when it receives the position of the cap by the **Interaction Scenario**.

3.4.2 Interaction Scenario

The interaction scenario consists of having the robot ask people to give it a baseball cap lying nearby on the floor (3 to 4 m away). We chose this scenario because it involves a short, voluntary, neutral and simple task (compared to engaging people to sell something for instance). Figure 3.5 illustrates the finite-state machine implemented in the **Interaction Scenario** module to manage how the robot interact with people in this context, with the dialogue generated by the robot presented in Table 3.1. The right part of Fig. 3.5 indicates which facial expressions the robot takes according to the states of the interaction scenario, which are :

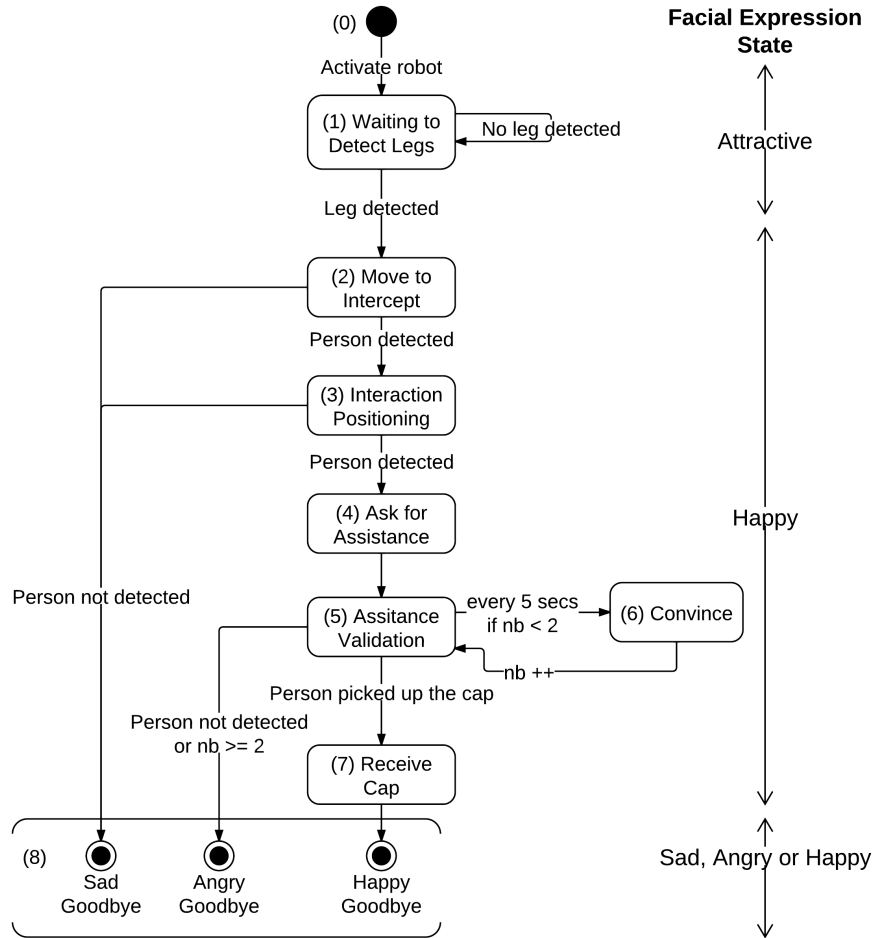


Figure 3.5 Finite-state machine of the **Interaction Scenario** module

- (0) This is an idle state during which the robot is waiting for the experimenter to initiate a trial. No behaviors are activated.

- (1) **Waiting to Detect Legs.** In this state, data is allowed to be processed by the **Leg Detection** module, in search of a potential person to interact with. No behaviors are activated.
- (2) **Move to Intercept.** Once the legs of a person are detected and tracked, the **Meet Person** behavior is activated until the **Body Shape Detection** returns a valid detection of a person, which means that the person is close by and in front of the robot. Data from the laser range finder are then inhibited (to minimize computational load) and body detection data are allowed to be sent to the **Attract Person** behavior.
- (3) **Interaction Positioning.** The **Attract Person** behavior is activated to try to keep an ideal distance between the robot and the person with whom the robot wants to interact. Figure 3.6 illustrates the trajectory resulting from states (0) to (3).
- (4) **Ask for Assistance.** If the person is located using the Kinect data, the **Show Object** (making the robot first try to direct the attention of the person toward the cap) and **Keep Engage** (to then turn back toward the person to keep the person engaged in the interaction) behaviors are activated.
- (5) **Assistance Validation.** In this state, the robot validates if the person is responding to its request for assistance. **Keep Engage** tries to keep visual contact with the person, and **Check Person** behavior verifies if the person has been detected in a $1\text{ m} \times 0.8\text{ m}$ range around the cap. If so, the **Interaction Scenario** module considers that the person has executed the robot's request.
- (6) **Convince.** If the person is still being detected but does not go take the cap, the robot vocally tries to convince the person to do so.
- (7) **Receive Cap.** In this state, the **Keep Engage** behavior is used to keep the person in the robot's field of view. Note that the robot does not detect that it receives the cap : it considers that the person has given the cap back after a certain time.
- (8) **Goodbye.** Depending on how ends the interaction with the person, vocal messages are generated and the **Disengage** behavior is activated to make the robot go back to its original position.

Table 3.2 shows the theoretical minimum and maximum amount of time each state can take, revealing that duration of an interaction ranges from 27 to 50 sec.

Table 3.1 Dialogue used for the interaction scenario

| ID | Dialogue in French (English) |
|-----|---|
| (2) | Bonjour, je voudrais te parler. (Hi! I would like to talk to you.) |
| (3) | J'ai quelque chose à te demander. (I have something to ask you.) |
| (4) | Peux-tu me rapporter ma casquette, s'il te plaît? (Can you bring me back my cap, please?) |
| (8) | Dommage ... au revoir. (Too bad ... goodbye.) |
| (6) | Allez, fais un effort! (Common, make an effort!) |
| (6) | Tu es sûr que tu ne veux pas m'aider. (Are you sure? You're sure you don't want to help me.) |
| (6) | Cela me ferait plaisir! (This would make me so happy!) |
| (7) | Merci beaucoup. (Thanks a lot.) |
| (8) | Bonne journée! (Have a good day!) |

Table 3.2 States duration of the interaction scenario

| States | Min (sec) | Max (sec) |
|--------|-----------|-----------|
| (0) | 0 | 0 |
| (1) | 5 | 5 |
| (2) | 4 | 10 |
| (3) | 4 | 10 |
| (4) | 3 | 3 |
| (5) | 3 | 4 |
| (6) | 0 | 10 |
| (7) | 3 | 3 |
| (8) | 5 | 5 |
| Total | 27 | 50 |

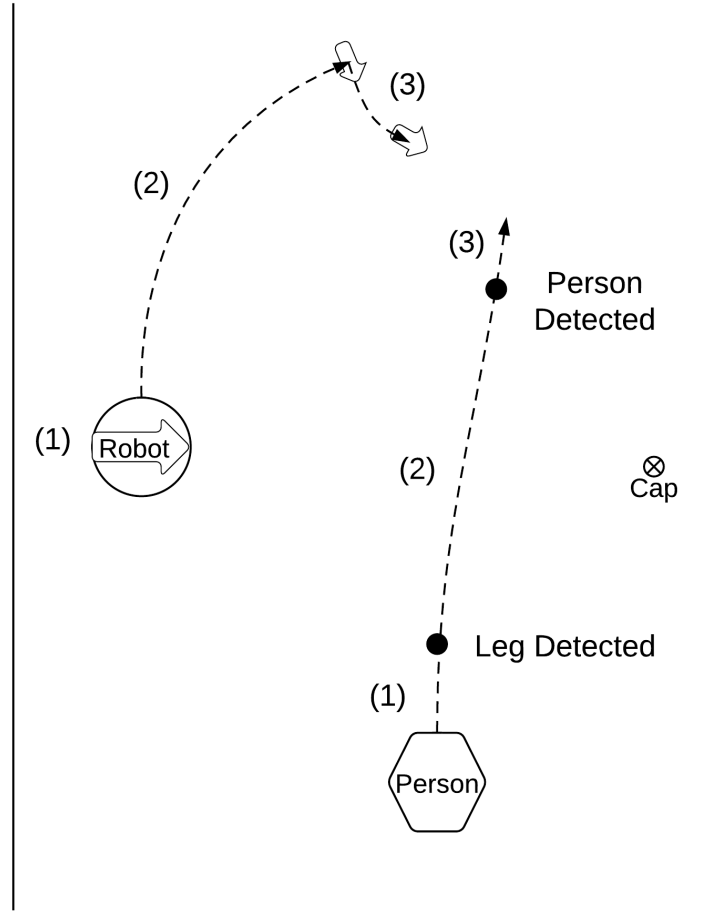


Figure 3.6 Trajectory of Johnny-Jr when approaching a person

3.5 HRI Evaluation with Johnny-Jr

Using the interaction scenario, the underlying objective of our case study is to examine the influences of verbal and non-verbal communication modalities in making a robot initiate interaction with people. By blocking commands to the actuators not involved in the intended test configuration, we made Johnny-Jr interact with people using the following sets of modalities :

- Voice and Facial expression modalities (VF). This is considered the baseline configuration, as it is usually what is used by interactive robots with facial expression, voice being the primary modality to provide basic information about the robot's intent.
- Voice, Facial expression and Head gesture modalities (VFH).
- Voice, Facial expression, Head gesture and Gesture modalities (VFHG).

- Voice, Facial expression, Head gesture, Gesture and Mobility modalities (VFHGM).
This configuration therefore integrates all of the interaction capabilities of Johnny-Jr.

Human-robot interaction studies bring the additional challenge of having to deal with the unpredictability of human responses. To acquire experience with different experimental frameworks, we decided to conduct a within-subject study [23] in controlled conditions, and a between-subject study [24] in open settings. For these trials, Johnny-Jr is always positioned according to Fig. 3.6 with the cap between three or four meters in front of the robot. The experiments received ethic approval with no use of video-recording (because some trials were done in public places).

3.5.1 Within-Subject Study in Controlled Conditions

The underlying objective of this experiment was to validate the implemented interaction scenario with people made aware that Johnny-Jr would engage interaction with them using different modalities. The test population consists of 35 able-bodied participants (23 male, 12 female), which can be characterized by people familiar with mobile robots (group A : 14 male, 1 female, all members of our lab), and people unfamiliar with mobile robots (group B : 9 male, 11 female). Each participant experimented all four interaction configurations (sequentially with about a minute pause between trials, selected in random order). The experimenter was located nearby without being too close to be considered to take part in the interaction. This allowed the experimenter to observe what is going on and to intervene for security purposes if necessary. Once done, they were asked to answer the following questions :

- Q1** Which configuration do you prefer ? Participants had to rank each configuration from 1 (most preferred) to 4 (least preferred). Our hypothesis is that people should prefer configurations with the highest number of interaction modalities, as suggested in Section 3.3.
- Q2** Which modality do you prefer ? Participants had to rank each communication modality from 1 to 5, to examine how the modality influences the preferences of the different configurations.
- Q3** How useful do you find each modality ? Participants had to indicate if the interaction modalities were found to be useful, useless or harmful to the interaction.

Participants were also allowed to provide general comments about their experience interacting with Johnny-Jr.

We first used the Freidman test [62], a non-parametric statistical tests for ranks of observations, to observe if a significant difference exists between configuration preferences (Q1) and between modality preferences (Q2). Results indicate that significant differences exist for Q1 ($p=0.000$) and for Q2 ($p=0.000$). We then used Equ. 3.1 and Equ. 3.2 [62] to identify the significant differences between the preference of each configuration for Q1 and of each modality for Q2. Equation 3.1 defines the minimum difference $\Delta\bar{R}_c$, with k being the number of compared conditions (four configurations for Q1, and five modalities for Q2), N the total number of participants (35), and $z_{\alpha/k(k-1)}$ the abscissa value from the unit normal distribution above which $\alpha/k(k-1)$ percent of the distribution lies (set according to $\alpha = 0.05$, the level of significance of the test) (ref. Appendix A of [62]). Equation 3.2 expresses that the averaged rank of preference of the two compared conditions (\bar{R}_u and \bar{R}_v) should exceed the minimum difference $\Delta\bar{R}_c$ to be considered to have significant differences. The average rank for each condition is calculated by adding every ranks of the condition and divided it by the number of participants in this condition. The resulting $\Delta\bar{R}_c$ values are 0.81 for Q1 and 1.06 for Q2.

$$\Delta\bar{R}_c = z_{\alpha/k(k-1)} \sqrt{\frac{k(k+1)}{6N}} \quad (3.1)$$

$$|\bar{R}_u - \bar{R}_v| \geq \Delta\bar{R}_c \quad (3.2)$$

Box plots are used to illustrate the results by representing outliers (circle), extremes (star), the smallest observation (lowest horizontal line), the lower quartile (lower end of the box), the median (horizontal line in the box), the upper quartile (upper end of the box) and the highest observation (highest horizontal line).

Preferred Configuration (Q1)

Fig. 3.7 illustrates for the test population, that the preferred configuration is VFHGM ($\bar{R}_{VFHGM} = 1.81$), followed by VFHG ($\bar{R}_{VFHG} = 2.01$), VFH ($\bar{R}_{VFH} = 2.64$) and finally VF ($\bar{R}_{VF} = 3.53$). Note however that significant differences are observed in Tab. 3.3 for all combinations between these configurations, except between VFHG - VFHGM and VFH - VFHG. This globally suggests that adding an interaction modality to the robot provides similar or improved interaction capabilities.

To analyze the results further, Fig. 3.8 illustrates the observed preferences for group A and group B. For each configuration, preferences for group A (left side) and for group B (right side) are represented. Preferences for group A are fairly consistent, as shown by the vertical lines and the segment : the preferences in order are VFGHM, VFHG, VFH and

VF. For group B, preferences are more scattered (as illustrated by the vertical boxes), and especially for the VFHGM configuration. Using the Mann-Whitney test [62] between preferences observed for group A and group B reveals that group A prefers VFHGM ($p = .043$) while group B prefers VFHG ($p = .086$). One possible explanation is that group B gets interested to the robot just from being in its presence, while group A is more interested when the robot interacts in more complex ways, such as when it moves.

Preferred Modality (Q2)

Fig. 3.9 illustrates modality preferences for the test population. The group clearly agrees about the rank of preferences for V ($\bar{R}_V = 1.64$, the highest) and F ($\bar{R}_F = 3.86$, the lowest), while for H ($\bar{R}_H = 3.04$), G ($\bar{R}_G = 2.89$) and M ($\bar{R}_M = 3.57$) the opinions are divided. Using Equ. 3.2 and Tab. 3.4, a significant difference is only observed between V and the other modalities, probably because the verbal modality is the most informative.

Looking at the preferences for A and B groups, Fig. 3.10 shows more differences in the evaluation provided by group A (compared to Q1), except regarding facial expressions (which comes last). For group B, preferences are also quite scattered. Using again the Mann-Whitney test, group A significantly prefers mobility compare to group B ($p=0.03$), while group B prefers the facial expression modality compared to group A ($p=0.019$). Other differences between groups A and B are not significant. This suggests that unfamiliar participants prefer interaction modalities that involve less motion from the robot.

Modality Usefulness (Q3)

Table 3.5 summarized the results obtained for Q3. Results suggest that voice is found to be useful by all (100%), and some participants commented that the robot's voice was clear and comprehensive. Gesture comes in second (77%), with 14% judging that the modality is useless and 9% finding it inappropriate. The interaction was found by some to be more user friendly when the robot was gesturing and pointing the baseball cap, but for others some gestures (Greeting/Goodbye, Come Closer) were difficult to understand. Head gesture is considered useless by 31% of the participants, but only 3% found in unsuitable. Some

Table 3.3 Difference ($|\bar{R}_u - \bar{R}_v|$) of average ranks of preferences between each configuration (Q1)

| | \bar{R}_{VF} | \bar{R}_{VFH} | \bar{R}_{VFHG} |
|-------------------|----------------|-----------------|------------------|
| \bar{R}_{VFH} | 0.89 | — | — |
| \bar{R}_{VFHG} | 1.51 | 0.63 | — |
| \bar{R}_{VFHGM} | 1.71 | 0.83 | 0.20 |

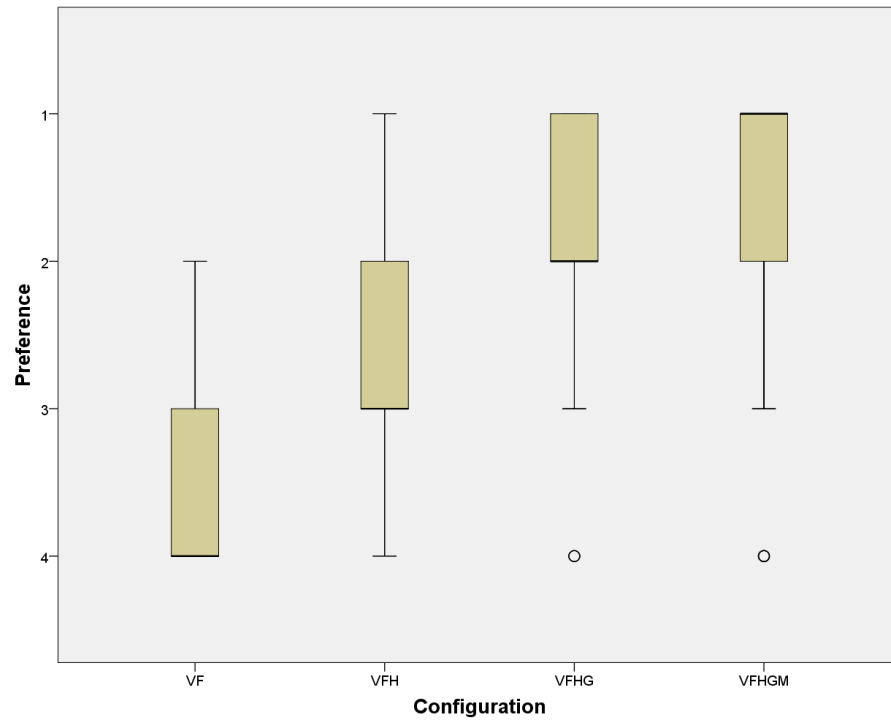


Figure 3.7 Q1 preferences for the test population

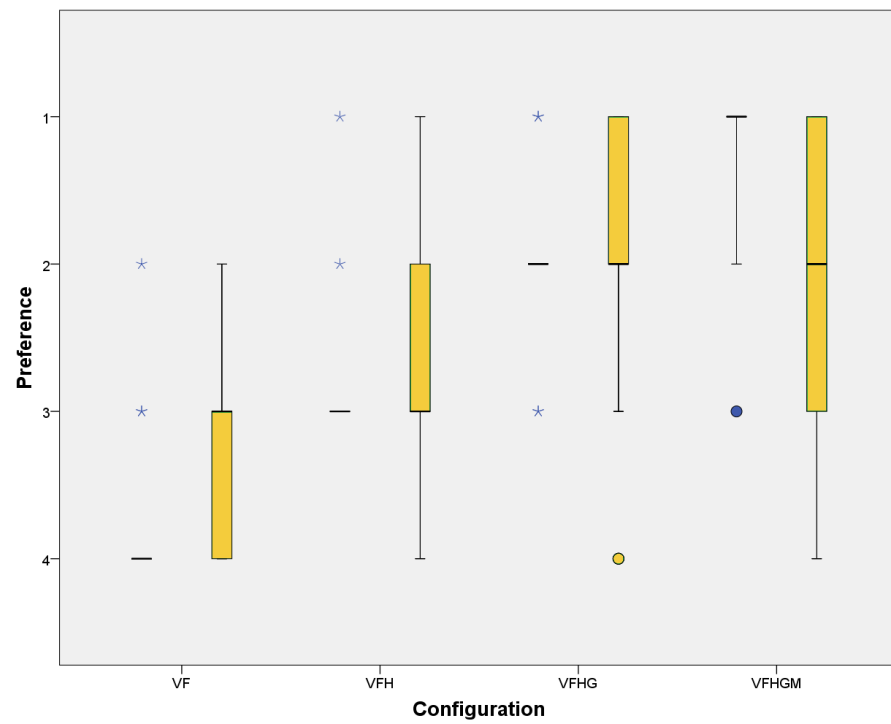


Figure 3.8 Q1 preferences for group A (left side - blue) and for group B (right side - yellow) for each configuration

Table 3.4 Difference ($|\bar{R}_u - \bar{R}_v|$) of average ranks of preferences between each modalities (Q2)

| | \bar{R}_V | \bar{R}_F | \bar{R}_H | \bar{R}_G |
|-------------|-------------|-------------|-------------|-------------|
| \bar{R}_F | 2.21 | — | — | — |
| \bar{R}_H | 1.40 | 0.81 | — | — |
| \bar{R}_G | 1.24 | 0.97 | 0.16 | — |
| \bar{R}_M | 1.93 | 0.29 | 0.53 | 0.69 |

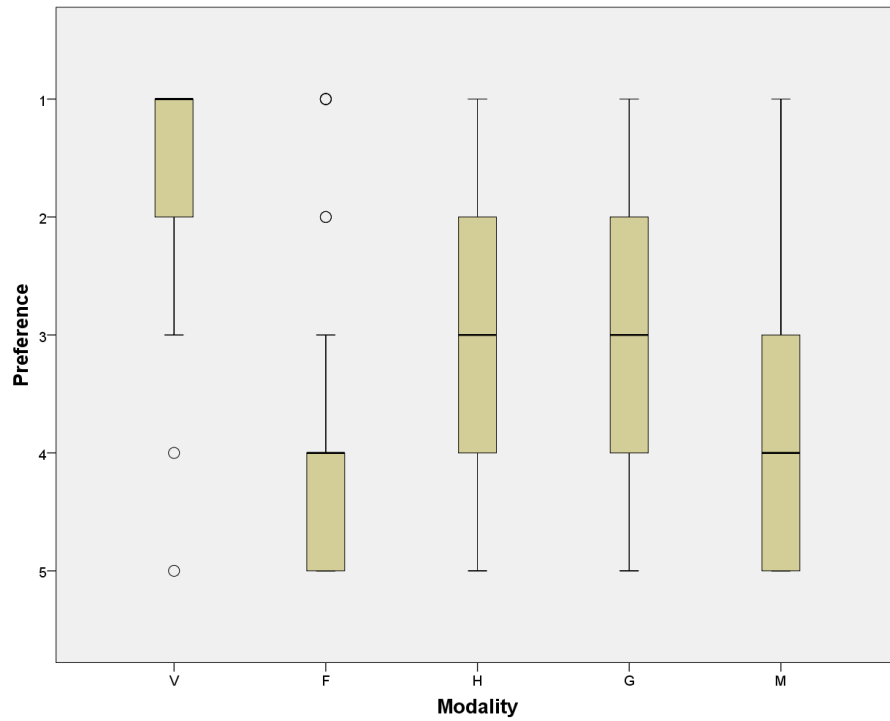


Figure 3.9 Q2 preferences for the test population

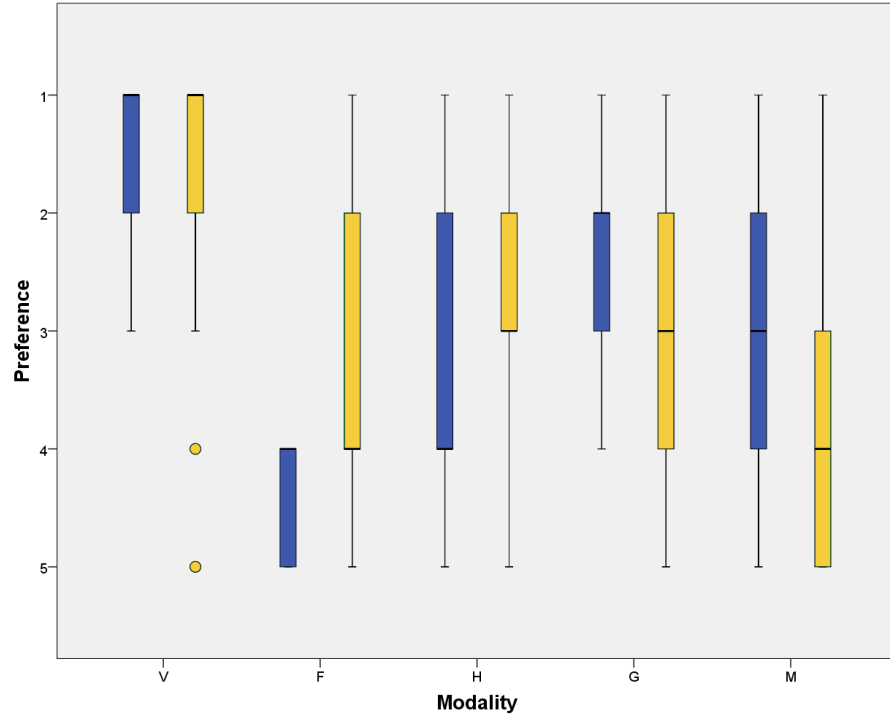


Figure 3.10 Q2 preferences for group A (left side - blue) and for group B (right side - yellow) for each modality

participants indicated that when the head was standing still, they did not feel that they were engaged in a conversation with the robot. One indicated that the pan-tilt unit was too slow and noisy. Mobility is considered useful by 60%, but useless by 17% and unsuitable by 23%. Three participants even felt fear as the robot moved toward them. This suggests that the trajectory made by the robot when approaching a person should be revised and improved. As examined in [17], making a robot approach a person from the front is less preferred compared to approach from the sides. Even though Johnny-Jr approach people from the side, it finishes by coming from the front to engage interaction.

Finally, about half of the participants found facial expression useful (and commented about the facial expressions were lovely and explicit), while the other half found it useless.

No significant difference between the groups A and B were found for Q3. The highest differences between the two groups are on the usefulness of gesture and of mobility. In those two cases, group B found the modalities less useful than group A, which suggests that gestures and mobility are not the first thing unfamiliar people notice on Johnny-Jr.

Table 3.5 Q3 results

| | Useful (A, B) | Useless (A, B) | Unsuitable (A, B) |
|-------------------|----------------------|-------------------|----------------------|
| Voice | 100% (100%, 100%) | 0% (0%, 0%) | 0% (0%, 0%) |
| Facial expression | 49% (40%, 55%) | 51% (60%, 45%) | 0% (0%, 0%) |
| Head gesture | 66% (73%, 60%) | 31% (27%, 35%) | 3% (0%, 5%) |
| Gesture | 77% (93%, 65%) | 14% (7%, 20%) | 9% (0%, 15%) |
| Mobility | 60% (80%, 45%) | 17% (7%, 25%) | 23% (13%, 30%) |

3.5.2 Between-Subject Study in Open Settings

The second set of trials was conducted over a two week period in five locations on our university campus, like the one shown in Fig. 3.11. According to Scopelliti et al. [57], a test population being mainly made of students is probably more receptive to robots and reacts more positively to it compare to the elderly people. Therefore, our results may not represent the perception of the population in general, but they are appropriate for the objectives pursued in our case study.

Participants were not made aware prior to interacting with the robot that it would try to engage interaction : they were solicited as they passed by the area where the robot was. To try to minimize the novelty factor that a robot may have when placed in a public space, we demonstrated Johnny-Jr's capabilities (not autonomous, using the remote gamepad controller) at these locations during the first week, to then experiment with the interaction scenario (autonomously executed by the robot) in the second week. We staid the entire day in the same location, testing one configuration in the morning and another in the afternoon.

Two experimenters were involved in these trials. They remained from 5 to 10 m away from the robot to intervene if needed while being able to observe without taking part in the interaction. One manually coded the reactions of participants had (e.g., start to look at the robot, look at the object shown, nod to the robot, speak, come closer, move away, stop or give back the cap) while interacting with the robot. The robot also memorized the duration of the interaction and the state at which participants left the interaction, and if they fulfilled the robot's request for assistance. Participants were then stopped



Figure 3.11 Johnny-Jr in position to interact with people in open settings, for a configuration not involving mobility (with the robot being connected to its charger)

by the second experimenter (who was also responsible to intervene using the gamepad if necessary), to see if they agreed to answer the following questions :

Q1 Is this your first encounter with the robot ?

Q2 Did you understand what the robot asked you ?

Q3 Did you find the robot's behavior coherent ?

Q4 Did you like the robot ?

Q5 Did you find the robot insistent ?

Q6 What is your confidence level for the robot ?

Q7 How much interacting with the robot seems similar to interacting with a human ?

Q8 What are the three primary modalities that made you stop to interact with the robot (Voice, Appearance, Face, Attitude (global behavior), Gesture (arm and head), Mobility) ?

The aim of these questions was to evaluate the quality of the interaction perceived by the participants for the different configurations. A Likert scale from 1 (lowest preference) to 5 (highest preference) was used for questions Q2 to Q7.

Overall, Johnny-Jr tried 1376 times to engage interaction (steps (1) and (2) of Fig. 3.5) with people passing by, with 381 trials (or 27.6%) having the robot asking for assistance. This is similar to the 22% reported in [14] about people stopping to interact with a robot in a museum. In our trials, this is explained by the fact that many people did not slow down to let the robot processed fast enough the perceived information, leaving Johnny-Jr's field of view before it could ask for assistance. This made Johnny-Jr either come back to its initial position, or continue interaction with another person detected in its field of view. Johnny-Jr was able to complete the interaction scenario with 102 participants (observing on average 5.5 reactions made by people while interacting with the robot, for interaction lasting on average 23.8 sec), and 99 participants accepted to fill out the questionnaire afterwards. For the other $381 - 102 = 279$ trials, the interaction failed because of limitations of Johnny-Jr in detecting or recognizing people in open and dynamic settings (e.g., one person would engage interaction and was replaced by another one passing in front of the robot; one person would interact with the robot while another grabbed the baseball cap).

Table 3.7 summarizes the results observed for the 99 valid trials. Unfortunately, we found no significant difference between the number of participants reactions, the interaction time or the number of complete evaluations compared to the number of trials. The average number of reactions ranges from 5 with VF to 6 with VFHGM : having longer interaction would probably help observe more significant differences. The ratio of number of valid trials to the number of trials demonstrate that VF has a lower ratio (5.7%) compare to configurations with more modalities (VFH : 9.1%, VFHG : 6.3% and VFHGM : 8.1%).

Table 3.7 also presents means and standard deviations observed Q1 to Q7. The results for Q1 indicate that more than half of the participants had already seen the robot. Participants rated the robot as very understandable (Q2). For Q3, the values for VFHG and VFHGM are slightly higher than the other configurations, and in general participants find the robot fairly coherent. For Q4, participants tend to less like the robot in the VF configuration comparing to the other configurations, but they highly appreciated the robot in general. The robot is considered less insistent in the VFHGM participant (Q5). From all the configurations, VF has the lower confidence level (Q6) for participants. For Q7, participants evaluate the interactions with the robot to have little resemblance to interactions with humans. Note however that no significant differences between these questions were

observed, except that participants that had previously seen the robot (Q1) found it to more insistent (Q5) ($p = .009$).

Table 3.6 summarizes what made participants stop to interact with Johnny-Jr (Q8). The percentages were calculated using the number of time the modality was chosen divided by the number of participants for each configuration. Preferences for participants who experienced VFHGM are Voice (78%), Appearance (52%), Mobility (43%) and Face (39%). As modalities are added to the interaction with the robot, we can see that Voice is similar in three configurations while Face decreases and Gesture and Mobility increase. But again, no significant differences were observed between these modalities. In general, most participants were surprised and excited to interact with a robot (which is unusual for them), so they liked whatever the robot did. In addition, participants interacted with Johnny-Jr using only one set of modalities, and they could not compare the influence of one modality over the others (as it was done in the within-subject study). Strangely, participants did indicate preferences for modalities that they did not experiment. Finally, some participants indicated that curiosity was the main factor that made them interact with Johnny-Jr.

Table 3.6 Results for Q8

| | Voice | Appearance | Face | Attitude | Gesture | Mobility |
|-------|-------|------------|------|----------|---------|----------|
| VF | 78% | 39% | 65% | 22% | 13% | 4% |
| VFH | 63% | 60% | 50% | 40% | 13% | 10% |
| VFHG | 78% | 52% | 52% | 39% | 22% | 0% |
| VFHGM | 78% | 52% | 39% | 22% | 30% | 43% |

Table 3.7 Data gathered with the 99 participants of the between-subject study in open settings

| Configurations | Q1 | Q2 | Q3 | Q4 | Q5 | Q6 | Q7 | Number of reactions | | Interaction time (sec) | Number of valid trials | Total number of trials |
|----------------|-----|------|------|------|------|------|------|---------------------|--|------------------------|------------------------|------------------------|
| VF : Mean | 52% | 4.61 | 3.65 | 3.65 | 2.35 | 2.89 | 2.41 | 5.04 | | 23.21 | 23 | 404 |
| Std. Err. | 11% | .12 | .30 | .26 | .20 | .25 | .19 | .30 | | .48 | | |
| VFH : Mean | 60% | 4.50 | 3.57 | 4.08 | 2.75 | 3.67 | 2.45 | 5.37 | | 25.29 | 30 | 325 |
| Std. Err. | 9% | .10 | .18 | .20 | .23 | .18 | .19 | .20 | | .82 | | |
| VFHG : Mean | 52% | 4.74 | 3.74 | 4.24 | 2.30 | 3.17 | 2.50 | 5.52 | | 22.86 | 23 | 364 |
| Std. Err. | 11% | .09 | .23 | .21 | .27 | .25 | .21 | .38 | | .60 | | |
| VFHGM : Mean | 48% | 4.43 | 3.76 | 4.13 | 2.17 | 3.30 | 2.61 | 5.96 | | 23.51 | 23 | 283 |
| Std. Err. | 11% | .13 | .20 | .21 | .24 | .28 | .27 | .34 | | .88 | | |

3.6 Discussion and Future Work

The objective of the case study conducted with Johnny-Jr was to go through a complete cycle of integrated design and evaluation of interaction modalities in natural and open settings. General recommendations that are being taken into consideration for the design of Johnny-0 are :

- Increased mobility may be suitable to create rich and engaging interaction with people, and may be affected by their familiarity with the robot. Johnny-0 has two compliant arms with five DOF, which should help generate clearer gestures. Special care must be taken when the robot approaches someone, and the impact of such capability should be validated specifically to pinpoint its effect on the interaction. Changing the pan-tilt unit of the head to minimize motor noise is also something that we are looking at. Saulnier et al. [56] observed that the speed of motion of the robot (mobile base and head) increases the interruption perception of the participant only when the robot is close to the participant. Slowing down the robot as it approaches people should be examined.
- People perception should be fast and accurate at the maximum possible range in order for the robot to engage interaction efficiently, because the robot has limited acceleration and speed.
- Participants were frequently seen responding back vocally to Johnny-Jr. This suggests that reciprocal interaction capabilities are expected by participants. Because Johnny-Jr interacted vocally with people, they expected the robot to be able to understand what they were saying. In that regard, Johnny-0 will be equipped with a sound source localization, tracking and separation system [74] with speech recognition capabilities. Recognition of facial expressions and of gestures would also be integrated on Johnny-0.
- Familiarity with the robot plays an important role in the evaluation of human-robot interaction, and influences the possibility of finding statistically significant results. Experimenting with people, and especially in open settings, is not easy : it requires time and effort, and detailed planning. At the same time, we believe that it is by conducting experiments 'in the wild' [54] that we can truly evaluate the influences of the interaction modalities of the platform. Trials with real people are quite demanding, and the experimental framework of this case study should be enhanced for the trials to be conducted with Johnny-0. For instance, within-subject studies in controlled conditions should be conducted to evaluate the effectiveness and

appropriateness of each interaction modalities individually, and then to integrate them to implement a specific interaction scenario. To minimize the curiosity bias, we would also recommend that before conducting a between-subject study in public places, we should make the robot interact with people at the experimental location over a long period of time, and use voluntary participants that are used to interact with the robot. For instance, we could have the robot advertise the lunch menu or important events at the cafeteria for a couple of days, to then experiment the interaction scenario with people that expect the robot to interact with them.

The design of Johnny-0 is well underway, and the observations made using Johnny-Jr reveals to be quite beneficial to guide the design and evaluation of its interaction capabilities. The study of interactive, autonomous humanoid robots is at its infancy, and lots of challenging interdisciplinary issues (involving engineering, programming, reasoning and human factor study) to tackle. By trying to establish common design and evaluation guidelines and frameworks, we believe that the case study reported in this paper can be beneficial to others as well.

3.7 Acknowledgements

This work was supported in part by the Natural Sciences and Engineering Research Council of Canada, the Canadian Foundation for Innovation and the Canada Research Chair program. The authors want to thank David Brodeur, Dominic Létourneau and all the volunteers who took part in the trials.

CHAPITRE 4

Démarche et contribution

Ce projet de maitrise a deux objectifs principaux. D'un côté, le développement et l'intégration de capteurs et d'actionneurs pour obtenir un robot humanoïde autonome avec un comportement naturel au sein d'un scénario défini. De l'autre, l'évaluation de l'interaction des personnes avec ce robot permet de donner des recommandations pour le design des prochains robots humanoïdes. Ce travail met ainsi en oeuvre le robot dans un scénario d'interaction autant en milieu ouvert qu'en laboratoire. Le but est de pouvoir en déduire la préférence d'interaction suivant la configuration du robot en jeu. Ceci est un travail préparatoire pour Johnny-0 afin de concevoir une interaction plus naturelle grâce aux conclusions de ce projet.

Le déroulement de ce projet de maitrise s'est fait en plusieurs phases :

- Définition de projet ;
- Développement des moyens de détection et des modalités d'interaction du robot, mise au point du protocole d'évaluation, développement des comportements du robot, mise en place du scénario d'interaction ;
- Évaluation en milieu ouvert et en laboratoire, interprétation des résultats et rédaction de l'article scientifique et du mémoire. Plusieurs pistes ont été explorées au niveau des moyens de détection et des algorithmes de traitement de données.

Tout d'abord, deux caméras et un laser ont été utilisés sur le robot. Une première caméra avec un faible angle de vue était fixée sur la tête du robot, couplée avec un algorithme de détection de visage [75]. Cela permettait au robot de maintenir l'attention d'une personne grâce au suivi de visage effectué par le pan-tilt. Celui-ci fait office de cou pour la tête de Johnny-Jr. La deuxième caméra omnidirectionnelle est fixée sur la base du robot, permettant au robot de visualiser la demi-sphère en face de lui. Un algorithme de détection de personne (HOG detector [16]) était ensuite appliqué à cette image, ce qui permettait au robot de savoir dans quelle direction une personne se trouvait. Les données générées par le laser Hokuyo ont été utilisées pour l'évitement d'obstacle et la détection de jambe [1]. Ce dernier algorithme a été adapté pour avoir moins de faux positifs en suivant une seule personne et en donnant sa direction de marche.

D'autres moyens de détections (un réseau de microphones et une Kinect) ont été intégrés pour améliorer la détection des personnes et ainsi la fluidité et le naturel du robot. Un réseau de huit microphones, exploité grâce à manyears, a été testé et intégré, mais il n'a finalement pas été utilisé, car peu pertinent dans le scénario choisi. En effet, les personnes étaient peu amenées à parler au robot. Ensuite, la Kinect a été introduite sur le robot. L'utilisation de l'API d'OpenNI a permis d'exploiter les données de la Kinect pour obtenir en temps réel la position courante du centre de masse de la personne détectée. La limitation de la puissance de calcul a été une barrière déterminante. En effet, il n'était pas possible de faire fonctionner tous les algorithmes de traitement de données en même temps.

Par conséquent, le choix s'est porté sur les deux moyens de détection les plus efficaces et complémentaires. En premier lieu, la détection grâce au laser Hokuyo qui donne une grande portée de détection et qui ajoute une information sur la direction de la personne en plus de sa position. En second lieu, la Kinect permet une détection réactive et précise avec peu de faux positifs comparés aux algorithmes utilisés avec les caméras standards.

Pour ce qui est des moyens d'interaction, ils ont été rajoutés au fur et à mesure du projet. Dans un premier temps, une base mobile avec deux roues motrices (Télérobot) a été utilisée. Cela a servi pour développer et intégrer la navigation, les expressions faciales, la synthèse vocale et les mouvements de la tête. La navigation du robot était basée sur le planificateur de trajectoire local de ROS qui génère une trajectoire en fonction des obstacles environnant grâce aux informations du laser et de l'odométrie. Ensuite, Johnny-Jr a réellement pris forme sur Azimut 3, la base mobile omnidirectionnelle d'IntRoLab. Le robot a ainsi gagné en rapidité et a pu générer des trajectoires plus complexes pour l'approche des personnes. Le contrôle via CIR (Centre Instantané de Rotation) de cette base a été possible grâce aux travaux de Lionel Clavier et François Ferland. Un seul bras articulé a finalement été ajouté pour générer les gestes. Les deux bras devant munir Johnny-0 n'étaient pas encore finalisés.

Quatre comportements ont été pensés pour mettre en place une interaction minimale avec une personne. Ces comportements sont l'approche, le maintien de l'attention, la présentation d'un objet et le désengagement. Chaque comportement du robot a été réalisé afin d'être indépendant et autonome des autres comportements. Le scénario d'interaction a été pensé pour que ce soit le robot qui engage la conversation avec des personnes pour les convaincre de faire une action. La contribution technique est présentée dans la figure 3.4 dans la section 3.4.1. La majorité des modules représentés dans la figure 3.4 ont été entièrement développés dans le cadre de cette maîtrise. Le module de détection des jambes

a été adapté comme précisé précédemment, tandis que les modules de la «Kinect», «Laser Range Finder» et «Base Mobile» proviennent de sources externes préalablement citées.

Au total cela a représenté 34200 lignes de code. Ce décompte ne tient pas en compte les trois modules que ce projet de maîtrise n'a pas développés, mais il contient les développements effectués sur télérobot (exploitation des différentes caméras et algorithmes de détection de personne). Ces lignes de codes peuvent être réparties en quatre catégories :

- Capteurs, actionneurs et algorithmes de traitement de données : 19400 lignes
- Comportement et arbitration : 5200 lignes
- Scénario d'interaction : 3500 lignes
- Configurations pour le lancement du robot et modèles pour la description mécanique (pour les robots Télérobot et Johnny-Jr) : 6100 lignes

Un protocole d'évaluation a été rédigé afin d'obtenir l'autorisation du Comité d'évaluation de la Faculté d'Éducation et Sciences Sociales (Annexe B). Ce document présente le robot, le scénario d'interaction et les moyens d'évaluation utilisés, le but étant de valider notre protocole d'évaluation.

Deux scénarios précis ont été testés. Dans un cas, le robot demande aux personnes si elles peuvent mettre une boule de papier dans la poubelle. Dans un autre cas, le robot demande à une personne si elle peut lui redonner sa casquette. Seul le deuxième scénario a été choisi pour l'expérimentation, car le geste demandé est moins naturel et demande une plus grande interaction de la personne avec le robot.

Le scénario a dans un premier temps été élaboré et testé en laboratoire puis optimisé en milieu ouvert (dans le couloir principal du bâtiment de la Faculté de Génie). C'est dans ces conditions que l'on a constaté que l'approche du robot n'était pas assez efficace. En effet, le robot ne parvenait pas à rattraper la personne et à rester dans son champ de vision. Par conséquent, la trajectoire du robot a été modifiée afin de décrire un arc de cercle dans la direction où se déplace la personne. Ce qui permet au robot d'être plus visible (le robot reste dans son champ de vision de la personne) et réactif, car les roues sont positionnées pour la trajectoire à effectuer.

Une fois le scénario finalisé, une démonstration a été faite pendant une semaine dans les différents lieux de l'Université de Sherbrooke où les évaluations étaient programmées (Faculté de Droit, d'Éducation, de Sciences et le Centre Sportif). Ces démonstrations avaient pour but d'accoutumer les personnes à la présence du robot pour réduire l'effet

de surprise. Le robot n'était pas autonome mais contrôlé à distance par un opérateur et il n'exécutait pas le scénario décrit précédemment.

La semaine suivant la démonstration, l'expérimentation se déroulait avec le robot effectuant le scénario prévu. Les interactions entre le robot et les personnes ont été analysées selon le protocole d'évaluation défini pour le milieu ouvert. Ce protocole consistait à alterner les configurations du robot afin de déterminer si les participants avaient une préférence parmi celles-ci. Le lieu d'expérimentation changeait toutes les demi-journées pour varier les personnes cotoyées par le robot, le scénario étant répétitif. Trois moyens d'évaluation ont été utilisés dans cette expérimentation. Deux évaluations sont faites durant l'interaction entre le robot et le participant (comptage des actions du participants et relevé des informations internes du robot sur les interactions lancées). Le troisième mode d'évaluation est un questionnaire rempli par les participants à la suite de l'interaction avec le robot. L'interprétation de ces résultats a été faite dans la section 3.5.2. Cette évaluation montre la complexité de l'évaluation des interactions en milieu ouvert.

Comme il était prévu par le protocole, une deuxième expérimentation en laboratoire a été mise en place étant donné que les résultats en milieu ouvert n'étaient pas suffisamment révélateur. Cette deuxième expérimentation utilise le même scénario que le premier. Par contre, les participants sont volontaires et conscients que le robot va interagir avec eux. De plus, les participants expérimentent aléatoirement les quatre configurations du robot. L'évaluation se fait ensuite uniquement sur l'appréciation des participants qui classent les configurations et modalités du robot suivant leur préférence. Des conclusions plus précises ont pu être déduites de cette deuxième expérimentation, tel que le montre l'analyse présentée dans la section 3.5.1. Le fait que les participants soient en mesure de comparer les quatre configurations du robot permet de mieux contraster leur préférence.

CHAPITRE 5

Conclusion

Telles que traitées dans le présent ouvrage, l'intégration et l'évaluation de capacités interactives sur un robot humanoïde représentent un défi technologique et expérimental important. Relever ces défis est toutefois essentiel afin d'arriver à concevoir des robots interactifs efficaces et appropriés pour être utilisés en situation réelle.

Les travaux présentés montrent l'intégration de modalités d'interaction verbale et non-verbales sur un robot humanoïde autonome, ainsi que l'évaluation de l'interaction autonome du robot avec des personnes dans un scénario précis en milieux contrôlé et non contrôlé.

Les résultats en milieu contrôlé montrent clairement l'influence des modalités sur la qualité de l'interaction du robot avec les participants. La tendance observée est que plus il y a de modalités utilisées par le robot, meilleure est l'interaction. Cependant, la modalité de déplacement divise la préférence des participants, en particulier entre les deux groupes (familiers et non-familiers avec la robotique). En effet, les personnes non-familières avec la robotique sont plus réservées sur cette modalité. Il faudra donc porter une attention particulière sur la modalité de déplacement dans les projets futurs.

Les résultats en milieu non-contrôlé ont démontré la difficulté inhérente d'arriver à obtenir des observations statistiquement significatives. Ceci peut s'expliquer en partie par l'élément de nouveauté qu'apporte la présence du robot dans un environnement public, le fait que les participants n'ont pas pu comparer les configurations entre elles, ou encore par la faible durée du scénario d'expérimentation. Cependant, l'expérimentation réalisée est l'une sinon la plus complète actuellement dans le domaine des interactions humain-robot avec une plateforme mobile et autonome.

La contribution scientifique de ce travail de maîtrise est réelle d'un côté sur la mise en oeuvre technique et de l'autre pour l'expérience de l'évaluation en milieu réel. En effet, Johnny-Jr, qui intègre cinq modalités d'interaction, démontre la faisabilité d'obtenir un robot mobile autonome pour initier et maintenir l'interaction avec des personnes. L'expérimentation en milieu ouvert a montré ses limites malgré le fait que trois moyens d'évaluation différents (données du robot, d'un observateur et du participant) ont été utilisés dans le

protocole. Mais cette première expérience en milieu ouvert est nécessaire pour donner un protocole de référence pour d'autres évaluations en situation réelle.

Enfin, les travaux réalisés servent de base pour la conception et l'évaluation de Johnny-0, un robot plus sophistiqué muni de deux bras pour des gestes plus complets, d'une caméra omnidirectionnelle pour une vision élargie de son environnement et une matrice de microphones pour détecter des sources de sons et reconnaître des personnes. La figure 5.1 présente une image de ce robot en cours de conception. Les résultats présentés dans ce mémoire alimentent les choix de conception pour Johnny-0, et permettront de réaliser des interactions plus complexes avec les personnes et sur de plus grandes périodes de temps.

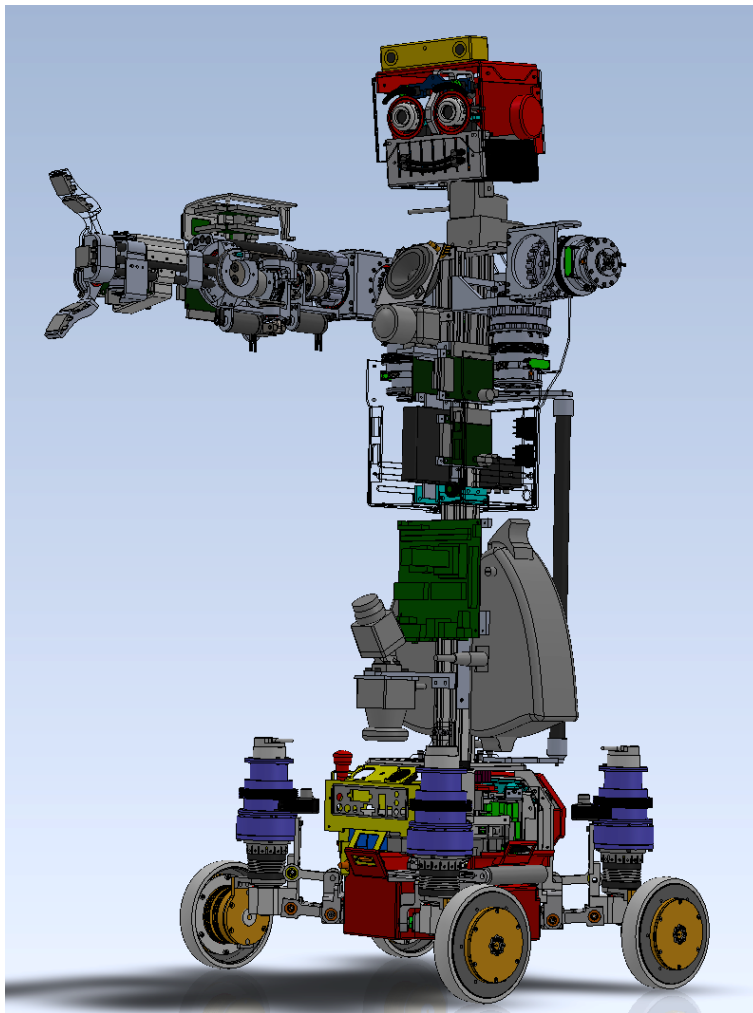


Figure 5.1 Illustration de Johnny-0

ANNEXE A



Figure A.1 Armar¹



Figure A.2 Asimo (Nakadai)²

¹source : <http://www.ira.uka.de/users/asfour/armar/armar.html>

²source : world.honda.com/news/2007/c071211Enabling-Multiple-ASIMO-to-Work/photo/index.html

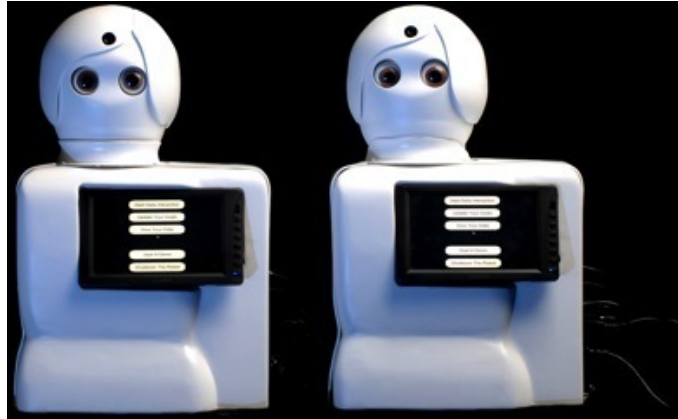


Figure A.3 Autom³



Figure A.4 Bandit II⁴

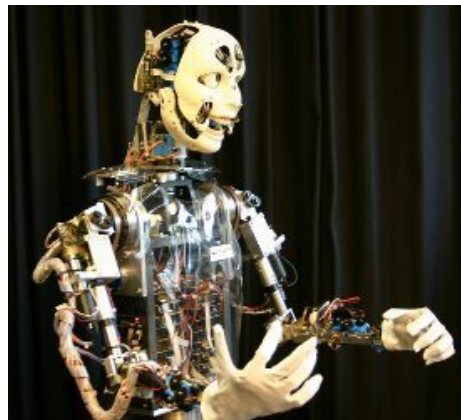


Figure A.5 Barthoc⁵

³source : <http://robotic.media.mit.edu/projects/robots/autom/overview/overview.html>

⁴source : <http://robotics.usc.edu/interaction/?l=Laboratory:Bandit2>

⁵source : <https://www.techfak.uni-bielefeld.de/ags/ai/projects/BARTHOC/welcome.html>

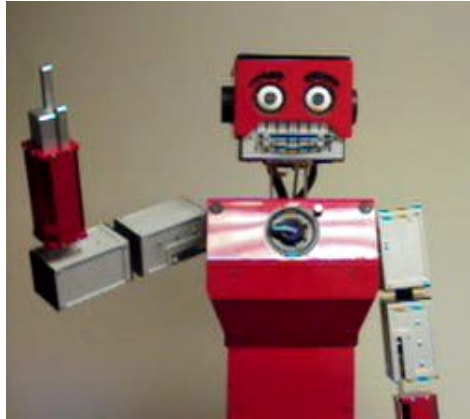


Figure A.6 Cramer (Scheutz)⁶



Figure A.7 Fritz⁷



Figure A.8 George⁸

⁶source : <http://cogs.indiana.edu/openhouse/>

⁷source : <http://nimbro.net/>

⁸source : article “Integrating vision and audition within a cognitive architecture to track conversations”
[71]



Figure A.9 GRACE⁹



Figure A.10 HRP-2W¹⁰

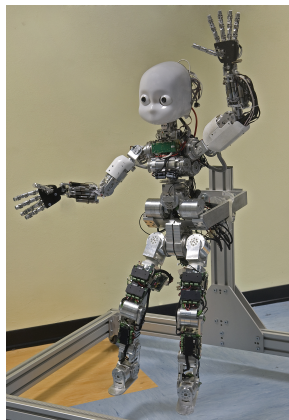


Figure A.11 iCub¹¹

⁹source : article “GRACE : an autonomous robot for the AAAI Robot challenge” [64]

¹⁰source : <http://www.jsk.t.u-tokyo.ac.jp/research/hrp2w/index.htm>

¹¹source : <http://eris.liralab.it/summerschool/>



Figure A.12 Justin¹²



Figure A.13 Kaspar¹³



Figure A.14 Keepon¹⁴

¹²source : http://www.dlr.de/rm/en/desktopdefault.aspx/tabid-5471/8991_read-16694/

¹³source : <http://kaspar.feis.herts.ac.uk/>

¹⁴source : <http://www.myu.ac.jp/xkozima/carebots/photo-eng.html>

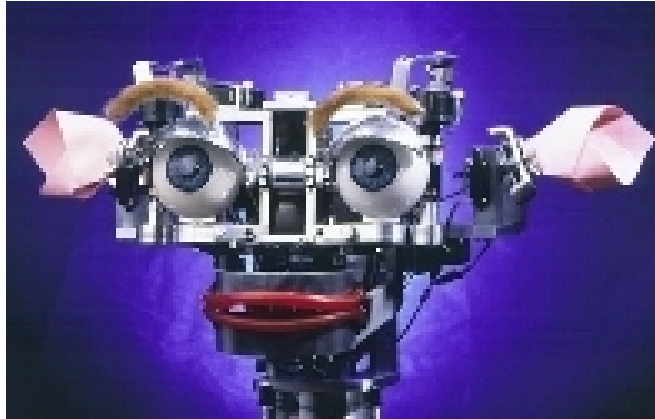


Figure A.15 Kismet¹⁵



Figure A.16 Leonardo¹⁶



Figure A.17 MDS¹⁷

¹⁵source : <http://www.ai.mit.edu/projects/sociable/overview.html>

¹⁶source : <http://robotic.media.mit.edu/projects/robots/leonardo/body/body.html>

¹⁷source : <http://www.nrl.navy.mil/aic/iss/index.php>



Figure A.18 Mel¹⁸



Figure A.19 Melvin¹⁹

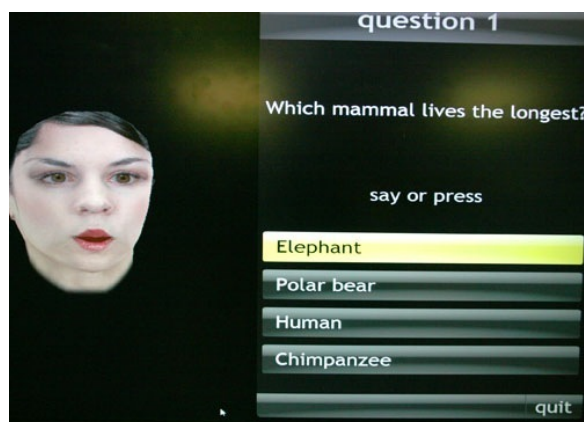


Figure A.20 Microsoft²⁰

¹⁸source : <http://www.merl.com/projects/hosting/>

¹⁹source : <http://web.cs.wpi.edu/~rich/hri/>

²⁰source : <http://www.techbubbles.com/microsoft/microsoft-research-projects/>

Figure A.21 PR2²¹Figure A.22 Rackham²²Figure A.23 Robotinho²³

²¹source : <http://www.willowgarage.com/pages/about-us/press-room>

²²source : <http://homepages.laas.fr/sara/laasko/>

²³source : <http://www.nimbro.net/robots.html>



Figure A.24 Robovie II ²⁴



Figure A.25 Robovie IV ²⁵

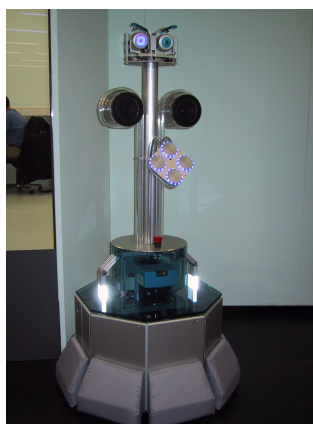


Figure A.26 RoboX²⁶

²⁴source : <http://www.is.sys.es.osaka-u.ac.jp/research/0006/index.en.html>

²⁵source : http://www.irc.atr.jp/en/research_project/humanoid/com_robo

²⁶source : http://projects.asl.ethz.ch/robox/database/exposition/just_robox/page_01.htm



Figure A.27 SIG²⁷



Figure A.28 Spartacus²⁸

²⁷source : <http://winnie.kuis.kyoto-u.ac.jp/SIG/>

²⁸source : <http://introlab.gel.usherbrooke.ca/mediawiki-introlab/index.php/Infrastructure>

ANNEXE B

Demande au comité d'éthique

Comité d'éthique de la recherche – Éducation et sciences sociales

Formulaire de demande d'évaluation éthique

1. INFORMATIONS GÉNÉRALES

Chercheuse ou chercheur principal (étudiante ou étudiant si c'est le cas) :

Nom : Vincent ROUSSEAU

Fonction : Étudiante à la maîtrise en génie électrique

Numéro de téléphone : 819-821-8000 poste 62312

Courriel : Vincent.Rousseau@USherbrooke.ca

Adresse postale : 181 rue Gillespie, Sherbrooke J1H 4X2

Équipe de direction dans le cas d'un projet d'étudiante ou d'étudiant : (*nom, fonction, numéro de téléphone, courriel, etc.*)

François Michaud, Professeur à l'Université de Sherbrooke et directeur de l'IntRoLab

Numéro de téléphone : 819-821-8000 poste 62107

Courriel : Francois.Michaud@USherbrooke.ca

Titre du projet : PROTOCOLE DE RECHERCHE POUR L'ÉVALUATION D'INTERACTION HUMAIN-ROBOT

Le projet est-il réalisé dans le cadre d'un programme d'étude?

Oui

Si oui, spécifier de quel programme il s'agit : Maîtrise en génie électrique

Le projet de recherche a-t-il été évalué et accepté par un jury? Non

Il n'y a pas de jury ou de comité d'éthique en faculté de Génie pour évaluer un projet touchant l'interaction humain-robot.

Est-ce que ce projet est subventionné?

Oui

Si oui, spécifiez l'organisme subventionnaire : CRSNG (Conseil de recherches en sciences naturelles et en génie du Canada)

2. Présentation du projet de recherche

Ce projet est un nouveau projet.

S'agit-il d'un projet interfacultaire ou interuniversitaire?

Non

À quelle date souhaitez-vous commencer la collecte de données?

Juin 2010 (1 mois)

3. Résumer le projet de recherche

Problématique

Le domaine de l'Interaction Humain-Robot (HRI) est en pleine expansion. En effet, de plus en plus de plateformes robotiques sont mises en œuvre pour faire évoluer ce domaine. Sur ces plateformes, toujours plus de modalités d'interaction sont mises en place telles que les mouvements corporels, la reconnaissance de geste ou d'objet, la reconnaissance et la synthèse vocale ou encore la mobilité, pour pouvoir effectuer l'interaction la plus complète et la plus naturelle pour l'humain. Mais ceci mène aussi à une complexité croissante de l'intégration de ces modalités sur une seule et même plateforme.

Aussi, le domaine HRI étant à ses débuts, la méthodologie expérimentale des travaux se limite le plus souvent à des preuves de concept éprouvées en laboratoire ou en milieux ouverts sans être bien contrôlées. Il se trouve que peu de chercheurs présentent une démarche structurée et rigoureuse pour l'évaluation, et il en résulte des recherches de types exploratoires qui examinent principalement la complexité technologique des modalités interactives à mettre en œuvre, et non l'impact de ces modalités sur la qualité des interactions. Ceci s'explique de par le fait que l'expertise des chercheurs en robotique porte sur la technologie, et non sur les expérimentations avec les humains. Cette observation s'applique aussi à nous, mais nous souhaitons aller plus loin avec nos travaux que la simple réalisation de modalités d'interaction sur un robot : nous voulons comprendre l'impact de ceux-ci lors d'interaction simples avec des humains.

Le but de l'étude est de mettre en valeur l'implication des gestes, de la vision, de l'audition, de la parole et la mobilité sur la qualité des interactions humain-robot.

Dans le domaine de l'interaction humain-robot (HRI, Human-Robot Interaction), il existe de multiples façons de dialoguer et d'interagir (par exemple : la parole, les gestes, le regard, la position du corps). Ce sont des choses que l'humain fait de façon intuitive et que le robot doit reproduire et s'adapter au mieux afin d'obtenir une communication intéressante et naturelle pour une personne.

Il s'avère donc intéressant de savoir quand et comment les modalités d'interaction du robot doivent être mises en jeu pour obtenir une communication adéquate avec l'humain. Pour cela, nous souhaitons étudier l'interaction entre le robot et l'humain dans le scénario suivant : *Le robot, placé dans un environnement ouvert (par exemple : corridor, grande salle), devra inciter les gens à poser une action (par exemple : entrer dans un local, apposer une signature, ramasser un objet).* Ce scénario simple permettra d'évaluer l'impact des différentes modalités soit une requête vocale, pointer l'objet, regarder l'objet, la direction du regard vers l'interlocuteur, se déplacer vers la personne pour l'inciter à interagir, et ce, avec des mesures simples comme le temps d'interaction, la réalisation de l'action ou encore l'appréciation des interactions. Le choix des conditions pour le type d'environnement ouvert et l'action à poser se fera en fonction des capacités technologiques en cours de développement.

Objectif principal

Déterminer la qualité de l'interaction entre l'humain et le robot suivant le type des modalités d'interaction mises en jeu. Ensuite, déterminer l'importance relative entre les modalités pour déterminer un ordre d'importance.

Hypothèses

La première hypothèse est que la qualité de l'interaction est différente suivant le type des modalités d'interaction mises en jeu.

La deuxième hypothèse est le classement par importance décroissante des modalités d'interaction:

- mobilité
- geste
- parole
- regard

Méthodologie

Le dispositif de recherche est un robot mobile humanoïde (figure 1) qui interagit avec les personnes évoluant autour de lui. Une personne est interpellée par le robot pour effectuer une action comme entrer dans un local, apposer une signature ou ramasser un objet. Le robot aura quatre modes de fonctionnement:

- Le robot reste au même endroit et interagit seulement avec la voix.
- Le robot utilise la voix et l'orientation du regard.
- Le robot utilise la voix, l'orientation du regard et les gestes.
- Le robot utilise la voix, l'orientation du regard, les gestes et le déplacement (selon la distance d'approche).

Le déplacement du robot pourra s'approcher d'une personne soit à une distance dite sociale (entre 120 et 300 cm) ou personnelle (entre 45 et 120 cm) [18]. Un biais possible serait un trop grand nombre de personnes autour du robot qui perturberait le comportement du robot et l'analyse de l'examineur. Pour éviter ce phénomène de curiosité, le robot sera laissé immobile sur les lieux pendant quelques jours avant l'expérimentation.

Participants

Les participants seront les personnes circulant à la Faculté de génie de l'Université de Sherbrooke. Ces personnes seront principalement des étudiants.

Procédures d'échantillonnage

La méthode d'échantillonnage sera un échantillonnage de commodité, car les participants dépendent de l'affluence du moment.

Procédures de collecte des données

La procédure de collecte des données sera effectuée tout d'abord par deux examinateurs. Le premier prendra en note les réactions de la personne aux différents stimuli du robot. Cet examinateur sera soit placé à portée de voix et de

vue de l'expérimentation, ou examinera la scène via une retransmission vidéo et sonore mais aucune donnée sera enregistrée. De plus, la volonté est que le participant ne remarque pas qu'il est observé pendant l'expérience. Le deuxième examinateur engagera la personne à remplir un questionnaire une fois l'interaction finie. Par la suite, un compte-rendu des actions du robot durant l'expérimentation sera fourni via un programme interne au robot.

Variables et instrument de mesure

Les variables suivantes seront mesurées grâce à l'observation et au compte-rendu des actions du robot :

- Temps de l'interaction.
- Nombres et types d'interaction entre l'humain et le robot (par exemple : parole, acquiescement, pointer, rapprochement, regard).
- Nombre de réponses aux stimuli du robot.
- Succès du robot à faire accomplir l'action par le participant.

Le sondage auprès des participants permettra de recueillir d'autres variables qualitatives :

- Amabilité du robot
- Clarté et compréhensibilité du robot
- Pertinence des actions du robot

Des tests préalables seront faits au laboratoire pour d'une part, régler le robot et ses comportements, et d'autre part, pour pouvoir entraîner l'observateur sur les informations à noter.

Taille de l'échantillon

Comme présenté précédemment, il y a quatre modes différents. L'objectif est d'avoir 20 participants par mode. Il est alors prévu de faire 80 observations avec des personnes différentes sur une période d'un mois.

Analyse des données

Le test statistique ANOVA (analyses de variance) sera utilisé pour analyser les données avec quatre modes différents et 20 échantillons par mode. Le but est de déterminer si les quatre modes génèrent des comportements différents. Pour analyser l'utilité de la modalité de déplacement vers la personne, le test t de Student est approprié, car il permet de vérifier si une différence significative existe entre deux échantillons indépendants. Le premier avec une distance d'approche sociale et le deuxième avec une distance d'approche personnelle plus oppressante.

Références :

IMAI , M., ONO , T., ISHIGURO , H. (Aug. 2003) Physical relation and expression : joint attention for human-robot interaction, IEEE Transactions on Industrial Electronics, vol. 50, no 4, p. 636-643.

SIDNER , C. L., LEE , C., MORENCY, L.-P., FORLINES , C. (2006) The effect of head-nod recognition in human-robot conversation, in Proceedings of the 1st ACM SIG-CHI/SIGART Conference on Human-Robot Interaction, p. 290-296, ACM, New York, NY, USA.

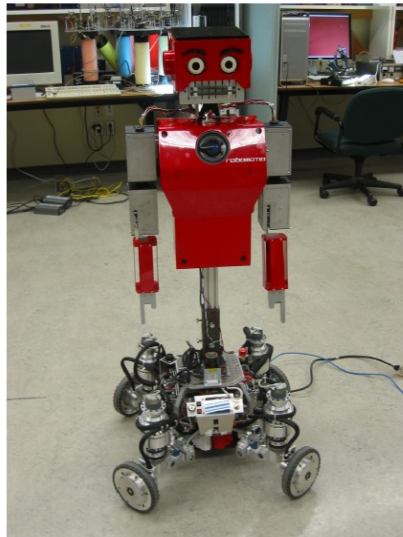


Figure 1 Image de la vue globale du robot Johnny-O

3. ASPECTS ÉTHIQUES

Équilibre entre les risques et les bénéfices

Quels sont les **risques encourus** par les personnes participantes?

Il n'y a aucun risque.

Plusieurs mesures vont être prises pour que les expériences se passent sans problème.

Tout d'abord, le robot ne fera pas de gestes brusques qui pourrait mettre en danger un participant.

De plus, le robot arrêtera tout déplacement lorsqu'une personne est proche. Enfin, un bouton d'arrêt d'urgence et l'envoi d'un signal d'arrêt à distance seront possibles. En soit, le robot est muni d'actionneurs élastiques, qui sont intrinsèquement sécuritaires en cas de contact physique entre le robot et le participant.

Combien de **temps** demandera la participation des personnes participantes?

Tout d'abord le temps d'interaction entre le robot et le participant sera inférieur à 5 minutes puis le questionnaire qui suivra sera de l'ordre de 5 minutes.

Quels sont les **bienfaits** qu'ils pourront retirer?

Ils vont participer à l'avancement des connaissances en robotique. Peut permettre de satisfaire une certaine curiosité envers les robots.

Est-ce qu'une **compensation**, monétaire ou autre, sera remise pour le temps, les déplacements, etc. occasionnés par la participation au projet? **Non**

Est-ce que ce projet de recherche se situe sous le **seuil de risque minimal**? **Oui**

Consentement libre et éclairé

Est-ce que la recherche est de nature **consensuelle**?

Non

Sinon, justifier :

En vue de tester les capacités d'interaction de notre robot et d'évaluer la pertinence des modalités d'interaction, il nous paraît nécessaire que les participants ne soient pas prévenus avant leur interaction avec le robot. En effet, un biais pourrait être introduit si les participants sont prévenus.

Dans le questionnaire qui suit l'interaction il sera expliqué au participant qu'il a été observé et qu'il lui appartient que l'on prenne en compte ou non les données que l'on a recueillies.

Quelles seront les **mesures prises pour s'assurer du consentement libre et éclairé des personnes participantes**?

Concrètement, comment seront recrutés les participants et participantes à la recherche?

Les participants seront les personnes circulant à la Faculté de génie de l'Université de Sherbrooke. Ces personnes seront principalement des étudiants.

Qui ira rencontrer les personnes participantes potentielles pour les inviter à participer au projet?

Suite à l'interaction entre le robot et le participant, un examinateur abordera le participant pour savoir s'il veut participer au projet de recherche en répondant au questionnaire, s'il refuse toute donnée recueillie sera effacée.

Est-ce que cette sollicitation se fera en personne par le chercheur ou la chercheuse, ou bien par quelqu'un d'autre (ex. : responsable de l'établissement, enseignant)?

La sollicitation se fera en personne par le chercheur.

Pour les personnes qui auront accepté de participer, est-ce qu'un formulaire de consentement sera signé?

Oui

Est-ce que la recherche implique des **personnes mineures ou légalement inaptes**?

Non

Confidentialité des données

Quelles seront les mesures prises pour assurer le **caractère confidentiel et anonyme des données**?

Utilisation de codes numériques ou de noms fictifs pour protéger l'identité des personnes

Aucun enregistrement audio ou vidéo ne sera fait de l'interaction avec le robot.

Où seront conservées les données? Seront-elles conservées sous clé? Les fichiers électroniques seront-ils protégés par un mot de passe?

Les fichiers seront conservés dans le laboratoire de robotique de l'Université de Sherbrooke à IntRoLab (local C1-5103), qui est fermé à clé. Les fichiers électroniques seront sur un ordinateur du laboratoire qui est protégé par un mot de passe.

Qui aura accès aux données? (Chercheuses et chercheurs, assistantes et assistants de recherche,...)

Ces personnes doivent être conscientes notamment à l'importance de respecter la confidentialité des données.

Le laboratoire est accessible seulement par les membres du laboratoire qui sont des étudiants et des professionnels de recherche. Les données sur l'ordinateur ne seront accessibles que par le chercheur responsable du projet.

Quand seront détruites les données brutes (questionnaires papier, cassettes d'entrevues, etc.)?

Les données brutes seront détruites au plus tard deux ans après la fin du projet.

Quels sont les **moyens de diffusion** prévus?

Les deux listes qui suivent présentent les cibles de conférences et de journaux où une publication d'article serait possible.

- Conférences

HRI Human Robot Interaction :

- Soumission : fin septembre
- Déroulement : début mars
- Lieu : Lausanne, Suisse en 2011

Humanoids IEEE-RAS/RSJ International Conference on Humanoid Robots :

- Soumission : début juillet
- Déroulement : début décembre
- Lieu : Lieu Paris, France en 2009

ICRA IEEE International Conference on Robotics and Automation :

- Soumission : mi-septembre
- Déroulement : début mai
- Lieu : Lieu Shanghai, Chine en 2011

- Journaux

- International Journal of Social Robotics
- International Journal of Humanoid Robotics
- Journal of Behavioral Robotics

Est-ce que les moyens de diffusion ou d'utilisation secondaire des données risquent de permettre l'identification des personnes ayant participé à la recherche?

Non

Est-ce qu'une **utilisation secondaire des données** est prévue?

Non

**En remettant votre demande au comité, vous vous engagez
à faire parvenir au comité un rapport annuel et un avis à la fin du projet.**

**Vous pouvez à cette fin remplir le formulaire de suivi disponible sur le site Internet du comité
d'éthique de la recherche Éducation et Sciences sociales et nous l'acheminer par courriel à l'adresse :
ethique.ess@USherbrooke.ca .**

**Également en cas de changements au projet,
vous devez aviser le comité à l'aide de ce même formulaire.**

Invitation à participer et formulaire de consentement pour le projet de recherche
PROTOCOLE DE RECHERCHE POUR L'ÉVALUATION D'INTERACTION
HUMAIN-ROBOT

Vincent Rousseau à la maîtrise en génie électrique au Laboratoire de Robotique Intelligente (IntRoLab)
Sous la direction de François Michaud, Professeur à l'Université de Sherbrooke et directeur de
l'IntRoLab.

Le chercheur responsable du projet a reçu un financement du CRSNG pour mener à bien ce projet de
recherche.

Madame,
Monsieur,

Nous vous invitons à participer à la recherche en titre. Ce projet de recherche a pour but d'étudier les interactions entre l'humain et le robot. Pour cela, nous mettons en action notre robot Johnny-0 dans nos milieux de vie et nous évaluons la réaction des personnes non préparées à interagir avec le robot. L'objectif est d'en déduire l'importance des modalités d'interaction (les gestes, le regard, la parole et la mobilité) du robot pour susciter l'intérêt et l'engagement d'un humain à interagir avec lui.

En quoi consiste la participation au projet?

Ce projet se base sur votre non préparation à interagir avec un robot. Ainsi, votre comportement est plus naturel pour que nous puissions analyser vos réactions et évaluer la qualité de l'interaction de notre robot. Pour cela, il se déroule en trois phases :

- Vous interagissez avec le robot et nous prenons en note vos réactions.
- Vous comprenez et signez le formulaire de consentement si vous le désirez, sinon nous ne tiendrons pas compte de l'observation qui vient d'être faite.
- Vous répondez au sondage (environ 5 minutes) lié à l'interaction avec le robot.

Le robot est situé dans nos milieux de vie à la Faculté de génie soit un couloir ou une salle commune.

Le seul inconvénient lié à votre participation est le temps consacré à la recherche, soit environ 5 minutes.

Qu'est-ce que le chercheur fera avec les données recueillies?

Pour éviter votre identification comme personne participante à cette recherche, les données recueillies par cette étude seront traitées de manière entièrement confidentielle. Les résultats de la recherche ne permettront pas d'identifier les personnes participantes. Les résultats seront diffusés dans le mémoire de maîtrise, dans des articles scientifiques et dans des congrès internationaux dans le domaine de la robotique. Les données recueillies seront conservées sous clé dans le laboratoire de robotique de l'Université de Sherbrooke et les seules personnes qui y auront accès sont le chercheur, son directeur de recherche et les membres du laboratoire (étudiants et professionnels de recherche). Les données seront détruites au plus tard en 2012 et ne seront pas utilisées à d'autres fins que celles décrites dans le présent document.

Est-il obligatoire de participer?

Non. La participation à cette étude se fait sur une base volontaire. Vous êtes entièrement **libre de participer ou non**, et de vous retirer en tout temps sans avoir à motiver votre décision ni à subir de préjudice de quelque nature que ce soit.

Y a-t-il des risques, inconvénients ou bénéfices?

Il n'y a aucun inconvénient direct pouvant découler de votre participation à ce projet. Sauf le temps requis pour remplir le questionnaire soit environ 5 minutes.

Aucune compensation d'ordre monétaire n'est accordée.

Que faire si j'ai des questions concernant le projet?

Si vous avez des questions concernant ce projet de recherche, n'hésitez pas à communiquer avec moi aux coordonnées indiquées ci-dessous.

[Signature de la chercheuse ou du chercheur] [Date de la signature]

Vincent Rousseau, étudiant à la maîtrise en génie électrique

Chercheur responsable du projet de recherche

Téléphone : 819-821-8000 poste 62312

Courriel : Vincent.Rousseau@USherbrooke.ca

*J'ai lu et compris le document d'information au sujet du projet **Protocole de recherche pour l'évaluation d'interaction humain-robot** J'ai compris les conditions, les risques et les bienfaits de ma participation. J'ai obtenu des réponses aux questions que je me posais au sujet de ce projet. J'accepte librement de participer à ce projet de recherche.*

Participante ou participant :

Signature :

Nom :

Date :

**S.V.P., signez les deux copies.
Conservez une copie et remettez l'autre au chercheur.**

Ce projet a été revu et approuvé par le comité d'éthique de la recherche Éducation et sciences sociales, de l'Université de Sherbrooke. Cette démarche vise à assurer la protection des participantes et participants.

Si vous avez des questions sur les aspects éthiques de ce projet (consentement à participer, confidentialité, etc.), n'hésitez pas à communiquer avec M. André Balleux, président de ce comité, au (819) 821-8000 poste 62439 ou à Andre.Balleux@USherbrooke.ca.

ANNEXE C

Questionnaire pour les expérimentations en milieu contrôlé

Sondage sur les modalités de Johnny-Jr

* Required

Sexe *

- ☐ Homme
☐ Femme

Préférence de configuration *

Classer les choix par ordre de préférence, '1' étant le meilleur et '4' le moins bon

| | 1 | 2 | 3 | 4 |
|---|-----------------------|-----------------------|-----------------------|-----------------------|
| Parole & émotion | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| Parole, émotion & mvt tête | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| Parole, émotion, mvt tête & geste | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| Parole, émotion, mvt tête, geste & mobilité | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |

Préférence de modalité dans l'interaction *

Classer les choix par ordre de préférence, '1' étant le meilleur et '5' le moins bon

| | 1 | 2 | 3 | 4 | 5 |
|----------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|
| Parole | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| Émotion | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| Mouvement de la tête | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| Geste | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| Mobilité | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |

Utilité de la modalité dans la compréhension du robot *

| | Utile | Inutile | Nuisible |
|----------------------|-----------------------|-----------------------|-----------------------|
| Parole | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| Émotion | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| Mouvement de la tête | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| Geste | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| Mobilité | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |

Commentaires :

Submit

ANNEXE D

Grille d'évaluation et questionnaire pour les expérimentations en milieu non-contrôlé

Sondage effectué après l'interaction avec Johnny-Jr

Date et Heure :

1 - Avez-vous compris ce que vous demandait le robot ?

Note de 1 à 5 (1 étant la moins bonne et 5 la meilleure) :

2 - Le comportement du robot était-il cohérent ?

Note de 1 à 5 (1 étant la moins bonne et 5 la meilleure) :

3 - Donnez votre degré d'affection pour le robot :

Note de 1 à 5 (1 étant nul et 5 le plus élevé) :

4 - Avez-vous trouvé le robot insistant ?

Note de 1 à 5 (1 étant peu et 5 beaucoup) :

5 - Donnez votre degré de confiance envers le robot :

Note de 1 à 5 (1 étant nul et 5 le plus élevé) :

6 - Qu'est ce qui chez le robot vous a donné envie d'interagir avec le robot ?

(3 choix max parmi les 6)

- | | |
|---|---|
| <input type="checkbox"/> Son apparence | <input type="checkbox"/> Ses gestes |
| <input type="checkbox"/> Son attitude | <input type="checkbox"/> Ses interventions vocales |
| <input type="checkbox"/> Ses déplacements | <input type="checkbox"/> Les mouvements de son visage |

7 - À quel point votre conversation avec le robot a-t-elle ressemblé à une conversation humain-humain?

Note de 1 à 5 (1 étant la moins bonne et 5 la meilleure) :

8 - Aviez vous déjà vu le robot ? ☐ oui ☐ non

LISTE DES RÉFÉRENCES

- [1] Arras, K. O., Grzonka, S., Luber, M. et Burgard, W. (2008). Efficient people tracking in laser range data using a multi-hypothesis leg-tracker with adaptive occlusion probabilities. Dans *Proceedings IEEE International Conference on Robotics and Automation*. p. 1710–1715.
- [2] Asfour, T., Regenstein, K., Azad, P., Schroder, J., Bierbaum, A., Vahrenkamp, N. et Dillmann, R. (2006). Armar-III : An integrated humanoid platform for sensory-motor control. Dans *Proceedings of 6th IEEE-RAS International Conference on Humanoid Robots*. p. 169–175.
- [3] Azad, P., Asfour, T. et Dillmann, R. (2007). Stereo-based 6D object localization for grasping with humanoid robot systems. Dans *Proceedings of IEEE/RSJ International Conference on Intelligent Robots and Systems*. p. 919–924.
- [4] Beaudry, E., Brosseau, Y., Carle, C., Raievsky, C., Létourneau, D., Kabanza, F. et Michaud, F. (2005). Reactive planning in a motivated behavioral architecture. Dans *Proceedings of the 20th National Conference on Artificial intelligence*.
- [5] Béchet, F. (2001). *Lia_phon : un système complet de phonétisation de textes, volume 42*. Traitement Automatique des Langues, p. 47–67.
- [6] Bennewitz, M., Faber, F., Joho, D. et Behnke, S. (2007). Fritz - A humanoid communication robot. Dans *Proceedings of the 16th IEEE International Symposium on Robot and Human interactive Communication*. p. 1072–1077.
- [7] Bohren, J., Rusu, R. B., Jones, E. G., Marder-Eppstein, E., Pantofaru, C., Wise, M., Mosenlechner, L., Meeussen, W. et Holzer, S. (2011). Towards autonomous robotic butlers : Lessons learned with the PR2. Dans *Proceedings IEEE International Conference on Robotics and Automation*.
- [8] Bohus, D. et Horvitz, E. (2009). Dialog in the open world : Platform and applications. Dans *Proceedings of International Conference on Multimodal Interfaces*.
- [9] Bohus, D. et Horvitz, E. (2009). Models for multiparty engagement in open-world dialog. Dans *Proceedings of the Workshop on Discourse and Dialogue*.
- [10] Breazeal, C. (2003). Toward sociable robots. *Robotics and Autonomous Systems*, volume 42, numéro 3-4, p. 167 – 75.
- [11] Breazeal, C., Kidd, C., Thomaz, A., Hoffman, G. et Berlin, M. (2005). Effects of non-verbal communication on efficiency and robustness in human-robot teamwork. Dans *Proceedings IEEE/RSJ International Conference on Intelligent Robots and Systems*. p. 708 – 713.
- [12] Breazeal, C. et Scassellati, B. (1999). How to build robots that make friends and influence people. Dans *Proceedings of International Conference on Intelligent Robots and Systems*. volume 2. p. 858–863.

- [13] Brick, T. et Scheutz, M. (2007). Incremental natural language processing for HRI. Dans *Proceedings of Second ACM/IEEE International Conference on Human-Robot Interaction*. p. 263–270.
- [14] Bruce, a., Nourbakhsh, I. et Simmons, R. (2002). The role of expressiveness and attention in human-robot interaction. Dans *Proceedings of the IEEE International Conference on Robotics and Automation*. p. 4138–4142.
- [15] Clodic, A., Fleury, S., Alami, R., Chatila, R., Bailly, G., Brethes, L., Cottret, M., Danes, P., Dollat, X., Elisei, F., Ferrane, I., Herrb, M., Infantes, G., Lemaire, C., Lerasle, F., Manhes, J., Marcoul, P., Menezes, P. et Montreuil, V. (2006). Rackham : An interactive robot-guide. Dans *Proceedings of 15th IEEE International Symposium on Robot and Human Interactive Communication*. p. 502–509.
- [16] Dalal, N. et Triggs, B. (2005). Histograms of Oriented Gradients for Human Detection. Dans Schmid, C., Soatto, S. et Tomasi, C., *International Conference on Computer Vision & Pattern Recognition (CVPR '05)*. volume 1. IEEE Computer Society, San Diego, États-Unis, p. 886–893.
- [17] Dautenhahn, K., Walters, M., Woods, S., Koay, K. L., Nehaniv, C. L., Sisbot, E. A., Alami, R. et Siméon, T. (2006). How may I serve you ? A robot companion approaching a seated person in a helping context. Dans *Proceedings of the ACM/IEEE International Conference on Human Robot Interaction*. p. 172–179.
- [18] DeVito, J., Chassé, G. et Vezeau, C. (2008). *La communication interpersonnelle*, 2^e édition. Éditions du Renouveau Pédagogique Inc. (ERPI).
- [19] Dutoit, T., Pagel, V., Pierret, N., Bataille, F. et van der Vrecken, O. (1996). The MBROLA project : Towards a set of high quality speech synthesizers free of use for non commercial purposes. Dans *Proceedings of Fourth International Conference on Spoken Language Processing*. p. 1393–1396.
- [20] Faber, F., Bennewitz, M., Eppner, C., Gorog, A., Gonsior, C., Joho, D., Schreiber, M. et Behnke, S. (2009). The humanoid museum tour guide Robotinho. Dans *Proceedings IEEE International Symposium on Robots and Human Interactive Communications*. p. 891–896.
- [21] Ferland, F., Clavien, L., Frémy, J., Létourneau, D., Michaud, F. et Lauria, M. (2010). Teleoperation of azimuth-3, an omnidirectional non-holonomic platform with steerable wheels. Dans *Proceedings IEEE/RSJ International Conference on Intelligent Robots and Systems*. p. 2515 –2516.
- [22] Groom, V., Takayama, L. et Nass, C. (2009). I am my robot : The impact of robot-building and robot form on operators. Dans *Proceedings of the Conference on Human-Robot Interaction*. p. 31–36.
- [23] Hall, R. (1998). *Within-Subjects Designs*. http://web.mst.edu/~psyworld/within_subjects.htm (page consultée le 22 septembre 2011).

- [24] Hall, R. (1998). *2x2 Between Subjects Factorial Design*. http://web.mst.edu/~psyworld/between_subjects.htm (page consultée le 22 septembre 2011).
- [25] Hoffman, G. et Breazeal, C. (2004). Robots that work in collaboration with people. Dans *Proceedings of the Conference on Human Factors in Computing Systems*.
- [26] Holroyd, A., Rich, C., Sidner, C. et Ponsler, B. (2011). Generating connection events for human-robot collaboration. Dans *Proceedings of the 6th ACM/IEEE International Conference on Human-Robot Interaction*.
- [27] Imai, M., Ono, T. et Ishiguro, H. (2003). Physical relation and expression : joint attention for human-robot interaction. *IEEE Transactions on Industrial Electronics*, volume 50, numéro 4, p. 636–643.
- [28] Inamura, T., Okada, K., Tokutsu, S., Hatao, N., Inaba, M. et Inoue, H. (2009). HRP-2W : A humanoid platform for research on support behavior in daily life environments. *Robotics and Autonomous Systems*, volume 57, numéro 2, p. 145–154.
- [29] Ishiguro, H., Ono, T., Imai, M., Maeda, T., Kanda, T. et Nakatsu, R. (2001). Robovie : An interactive humanoid robot. *An International Journal on Industrial Robot*, volume 28, p. 498–504.
- [30] Kanda, T., Glas, D., Shiomi, M. et Hagita, N. (2009). Abstracting people’s trajectories for social robots to proactively approach customers. *IEEE Transactions on Robotics*, volume 25, numéro 6, p. 1382–1396.
- [31] Kidd, C. et Breazeal, C. (2008). Robots at home : Understanding long-term human-robot interaction. Dans *Proceedings of IEEE/RSJ International Conference on Intelligent Robots and Systems*.
- [32] Kose-Bagci, H., Dautenhahn, K. et Nehaniv, C. L. (2008). Drumming with a humanoid robot : Results from human-robot interaction studies. Dans *Proceedings of ECSIS Symposium on Learning and Adaptive Behaviors for Robotic Systems*.
- [33] Kose-Bagci, H., Dautenhahn, K., Syrdal, D. S. et Nehaniv, C. L. (2007). Drummate : A human-humanoid drumming experience. Dans *Proceedings of IEEE-RAS Humanoids*.
- [34] Kozima, H., Michalowski, M. P. et Nakagawa, C. (2008). Keepon : A playful robot for research, therapy, and entertainment. *International Journal of Social Robotics*, volume 1, p. 3–18.
- [35] Lauria, M., Legault, M.-A., Lavoie, M.-A. et Michaud, F. (2008). Differential elastic actuator for robotic interaction tasks. Dans *Proceedings IEEE International Conference on Robotics and Automation*. p. 3606 –3611.
- [36] Lauria, M., Michaud, F., Legault, M.-A., Létourneau, D., Rétornaz, P., Nadeau, I., Lepage, P., Morin, Y., Gagnon, F., Giguère, P., Frémy, J. et Clavien, L. (2008). Elastic locomotion of a four steered mobile robot. Dans *Proceedings IEEE/RSJ International Conference on Intelligent Robots and Systems*. p. 2721 –2722.

- [37] Lee, J. K. et Breazeal, C. (2010). Human social response toward humanoid robot's head and facial features. Dans *Proceedings of the Conference on Human Factors in Computing Systems*. p. 4237.
- [38] Legault, M.-A., Lavoie, M.-A., Cabana, F., Jacob-Goudreau, P., Letourneau, D., Michaud, F. et Lauria, M. (2008). Admittance control of a human centered 3 DOF robotic arm using differential elastic actuators. Dans *Proceedings IEEE/RSJ International Conference on Intelligent Robots and Systems*. p. 4143–4144.
- [39] Michalowski, M. P., Sabanovic, S., DiSalvo, C. F., Font, D. B., Hiatt, L. M., Melchior, N. et Simmons, R. (2007). Socially distributed perception : Grace plays social tag at aaai 2005. *Autonomous Robots*, volume 22, numéro 4, p. 385–397.
- [40] Michaud, F., Audet, J., Létourneau, D., Lussier, L., Théberge-Turmel, C. et Caron, S. (2001). Experiences with an autonomous robot attending AAAI. *IEEE Intelligent Systems*, volume 16, numéro 5, p. 23–29.
- [41] Michaud, F., Côté, C., Létourneau, D., Brosseau, Y., Valin, J. M., Beaudry, E., Raïevsky, C., Ponchon, A., Moisan, P., Lepage, P., Morin, Y., Gagnon, F., Giguère, P., Roux, M. A., Caron, S., Frenette, P. et Kabanza, F. (2007). Spartacus attending the 2005 AAAI conference. *Autonomous Robots*, volume 22, numéro 4, p. 369–383.
- [42] Michaud, F., Ferland, F., Létourneau, D., Legault, M.-A. et Lauria, M. (2010). Toward autonomous, compliant, omnidirectional humanoid robots for natural interaction in real-life settings. *Paladyn Behavioral Robotic Journal*, volume 1, numéro 1, p. 57–65.
- [43] Michaud, F., Létourneau, D., Fréchette, M., Beaudry, E. et Kabanza, F. (2006). Spartacus, scientific robot reporter. Dans *Proceedings of the AAAI Mobile Robot Workshop*.
- [44] Mitsunaga, N., Miyashita, T., Ishiguro, H., Kogure, K. et Hagita, N. (2006). Robovie-IV : A communication robot interacting with people daily in an office. Dans *Proceedings of IEEE/RSJ International Conference on Intelligent Robots and Systems*. p. 5066–5072.
- [45] Mitsunaga, N., Miyashita, Z., Shinozawa, K., Miyashita, T., Ishiguro, H. et Hagita, N. (2008). What makes people accept a robot in a social environment, discussion from six-week study in an office. Dans *Proceedings of International Conference on Intelligent Robots and Systems*. p. 3336–3343.
- [46] Morency, L.-P., Sidner, C., Lee, C. et Darrell, T. (2005). Contextual recognition of head gestures. Dans *Proceedings of 7th International Conference on Multimodal Interactions*. p. 18–24.
- [47] Nakadai, K., Yamamoto, S., Okuno, H., Nakajima, H., Hasegawa, Y. et Tsujino, H. (2008). A robot referee for rock-paper-scissors sound games. Dans *Proceedings of IEEE International Conference on Robotics and Automation*. p. 3469–3474.
- [48] Okuno, H., Nakadai, K., Hidai, K., Mizoguchi, H. et Kitano, H. (2001). Human-robot interaction through real-time auditory and visual multiple-talker tracking. Dans *Pro-*

- ceedings of IEEE/RSJ International Conference on Intelligent Robots and Systems.* volume 3. p. 1402–1409 vol.3.
- [49] Okuno, H. G., Nakadai, K. et Kitano, H. (2002). Social interaction of humanoid robotbased on audio-visual tracking. Dans *Proceedings of the 15th International Conference on Industrial and Engineering Applications of Artificial Intelligence and Expert Systems*. Springer-Verlag, London, UK, p. 725–735.
- [50] Ott, C., Eiberger, O., Friedl, W., Bauml, B., Hillenbrand, U., Borst, C., Albu-Schaffer, A., Brunner, B., Hirschmuller, H., Kielhofer, S., Konietzschke, R., Suppa, M., Wimbock, T., Zacharias, F. et Hirzinger, G. (2006). A humanoid two-arm system for dexterous manipulation. Dans *Proceedings of 6th IEEE-RAS International Conference on Humanoid Robots*. p. 276–283.
- [51] Quigley, M., Gerkey, B., Conley, K., Faust, J., Foote, T., Leibs, J., Berger, E., Wheeler, R. et Ng, A. (2009). ROS : an open-source Robot Operating System. Dans *Proceedings IEEE International Conference on Robotics and Automation, Workshop on Open Source Software*.
- [52] Rich, C., Ponsler, B., Holroyd, A. et Sidner, C. L. (2009). Recognizing engagement in human-robot interaction. Dans *Proceedings of the Conference on Human Robot Interaction*.
- [53] Rich, C., Sidner, C. L. et Lesh, N. (2001). Collagen : Applying collaborative discourse theory to human-computer interaction. *AI Magazine*, volume 22, numéro 4, p. 15–25.
- [54] Salter, T., Michaud, F. et Larouche, H. (2010). How wild is wild ? A taxonomy to categorize the wildness of child-robot interaction. *International Journal of Social Robotics*, volume 2, numéro 4, p. 405–415.
- [55] Satake, S., Kanda, T., Glas, D. F., Imai, M., Ishiguro, H. et Hagita, N. (2009). How to approach humans ? : Strategies for social robots to initiate interaction. Dans *Proceedings of the 4th International Conference on Human Robot Interaction*. ACM, New York, NY, USA, p. 109–116.
- [56] Saulnier, P. (2011). *Exploring Socially Appropriate Nonverbal Robotic Interruption*. Mémoire de maîtrise, University of Calgary.
- [57] Scopelliti, M., Giuliani, M. V. et Fornara, F. (2005). Robots in a domestic setting : A psychological approach. *Universal Access in the Information Society*, volume 4, numéro 2, p. 146–155.
- [58] Sidner, C. L., Kidd, C. D., Lee, C. et Lesh, N. (2004). Where to look : A study of human-robot engagement. Dans *Proceedings of International Conference on Intelligent User Interfaces*.
- [59] Sidner, C. L., Lee, C., Kidd, C. D., Lesh, N. et Rich, C. (2005). Explorations in engagement for humans and robots. *Journal Artificial Intelligence*, volume 166, numéro 1-2, p. 140–164.

- [60] Sidner, C. L., Lee, C. et Lesh, N. (2003). The role of dialog in human robot interaction. Dans *Proceedings of First International Workshop on Language Understanding and Agents for Real World Interaction*.
- [61] Sidner, C. L., Lee, C., Morency, L.-P. et Forlines, C. (2006). The effect of head-nod recognition in human-robot conversation. Dans *Proceedings of the 1st ACM SIGCHI/SIGART Conference on Human-Robot Interaction*. ACM, New York, NY, USA, p. 290–296.
- [62] Siegel, S. (1988). *Nonparametric Statistics for the Behavioral Sciences*, 2^e édition. New York : McGraw-Hill, 399 p. p.
- [63] Siegwart, R., Arras, K. O., Bouabdallah, S., Burnier, D., Froidevaux, G., Greppin, X., Jensen, B., Lorotte, A., Mayor, L., Meisser, M., Philippsen, R., Piguet, R., Ramel, G., Terrien, G. et Tomatis, N. (2003). Robox at expo.02 : A large scale installation of personal robots. *Robotics and Autonomous Systems, Special issue on Socially Interactive Robots*, volume 42, p. 203–222.
- [64] Simmons, R., Goldberg, D., Goode, A., Montemerlo, M., Roy, N., Sellner, B., Urmson, C., Schultz, A., Abramson, M., Adams, W., Atrash, A., Bugajska, M., Coblenz, M., MacMahon, M., Perzanowski, D., Horswill, I., Zubek, R., Kortenkamp, D., Wolfe, B., Milam, T. et Maxwell, B. (2003). Grace : An autonomous robot for the AAAI robot challenge. *AI Magazine*, volume 24, numéro 2, p. 51–72.
- [65] Spexard, T., Haasch, A., Fritsch, J. et Sagerer, G. (2006). Human-like person tracking with an anthropomorphic robot. Dans *Proceedings of IEEE International Conference on Robotics and Automation*. p. 1286–1292.
- [66] Steinfeld, A., Fong, T., Kaber, D., Lewis, M., Scholtz, J., Schultz, A. et Goodrich, M. (2006). Common metrics for human-robot interaction. Dans *Proceedings of the 1st ACM SIGCHI/SIGART Conference on Human Robot Interaction*. ACM, New York, NY, USA, p. 33–40.
- [67] Takayama, L. et Pantofaru, C. (2009). Influences on proxemic behaviors in human-robot interaction. Dans *Proceedings IEEE/RSJ International Conference on Intelligent Robots and Systems*. p. 5495–5502.
- [68] Tapus, A., Fasola, J. et Mataric, M. J. (2008). Cognitive assistance and physical therapy for dementia patients using socially assistive robots. Dans *Proceedings of Social Interaction with Intelligent Indoor Robots*.
- [69] Taylor, P., Black, A. W. et Caley, R. (1998). The architecture of the Festival speech synthesis system. Dans *Proceedings of The Third ESCA Workshop in Speech Synthesis*.
- [70] Trafton, J., Fransen, B., Harrison, A. et Bugajska, M. (2009). An embodied model of infant gaze-following. Dans Naval Research Lab., Washington, D., *Technical Report*.
- [71] Trafton, J. G., Bugajska, M. D., Fransen, B. R. et Ratwani, R. M. (2008). Integrating vision and audition within a cognitive architecture to track conversations.

- Dans *Proceedings of the 3rd ACM/IEEE International Conference on Human Robot Interaction*. ACM, New York, NY, USA, p. 201–208.
- [72] Tsakarakis, N., Metta, G., Sandini, G., Vernon, D., Beira, R., Becchi, F., Righetti, L., Santos-Victor, J., Ijspeert, A., Carrozza, M. et Caldwell, D. (2007). iCub - The design and realization of an open humanoid platform for cognitive and neuroscience research. *Journal of Advanced Robotics, Special Issue on Robotic platforms for Research in Neuroscience*, volume 21, numéro 10, p. 1151–1175.
- [73] Tsui, K. T. et Yanco, H. A. (2009). Towards establishing clinical credibility for rehabilitation and assistive robots through experimental design. Dans *Proceedings Robotics : Science and Systems Workshop on Good Experimental Methodology in Robotics*.
- [74] Valin, J.-M., Michaud, F. et Rouat, J. (2007). Robust localization and tracking of simultaneous moving sound sources using beamforming and particle filtering. *Robotics and Autonomous Systems*, volume 55, numéro 3, p. 216–228.
- [75] Viola, P. et Jones, M. J. (2004). Robust real-time face detection. *International Journal of Computer Vision*, volume 57, numéro 2, p. 137–154.
- [76] Wainer, J., Feil-Seifer, D. J., Shell, D. A. et Mataric, M. J. (2006). The role of physical embodiment in human-robot interaction. Dans *Proceedings of the International Workshop on Robot and Human Interactive Communication*. p. 117–122.
- [77] Walker, W., Lamere, P., Kwok, P., Raj, B., Singh, R., Gouvea, E., Wolf, P. et Woelfel, J. (2004). *Sphinx-4 : A Flexible Open Source Framework for Speech Recognition* (Rapport technique).
- [78] Yamaoka, F., Kanda, T., Ishiguro, H. et Hagita, N. (2006). How contingent should a communication robot be? Dans *Proceedings of the 1st ACM SIGCHI/SIGART Conference on Human-Robot Interaction*. ACM, New York, NY, USA, p. 313–320.

